Exploration of the Knowledge of and Motivation for Learning Preservation Practices for Personal Digital Information

A Thesis

Submitted to the Faculty

of

Drexel University

by

Andrea Copeland Japzon

In partial fulfillment of the

Requirements for the degree

of

Doctor of Philosophy

June 2009

Dedications

To Denise, who for four years mentored me with a generous heart and a brilliant mind.

And to Jeff and Rocky for all the love.

Acknowledgements

Because the dissertation became my life, many people in my life became instrumental to its completion. I offer my biggest thanks and appreciation to them.

My kind and keenly insightful committee: Denise Agosto (Chair), Eileen Abels, Deborah Barreau, Greg Hislop, and Lisl Zach.

The three other members of the fab four Joan Beaudoin, Ann Bui, and Jennifer Burke who helped me continually with just about everything. Mike Atwood is right –much better to be a titmouse than a robin.

Four friends from the great state of Maryland who provided an endless supply of love and encouragement: Bridget Burns, Kathleen McGlaughlin, Nadia Nasr, and Summer Rosswog.

And four friends from Brooklyn: James Diamond, Robyn Kopet, Laura Kortz, and Matthew Wills, who helped me down the homestretch with great food, sound advice, spaces to work, and help when I needed it most.

Xia Lin and the Institute of Museum and Library Services for investing in me.

Mom and Dad for believing in me.

Table of Contents

List of Tables

List of Figures

Abstract

Exploration of the Knowledge of and Motivation for Learning Preservation Practices for Personal Digital Information

Andrea Japzon

Denise E. Agosto, Ph.D., Dissertation Advisor

This research investigated preservation practices related to personal digital information. It sought to answer three main research questions: 1. What kinds of personal digital information do public library users collect and why? 2. What are the cognitive, affective, and psycho-social influences that contribute to the preservation of personal digital information? 3. How can individuals improve their digital preservation practices and what would motivate them to make improvements to these practices?

Twenty-six participants were recruited from public library friends groups from a large metropolitan region on the east coast of the United States. Three areas within the study of information behavior informed the research design: everyday life information seeking (ELIS); information horizons and information source horizons; and cognitive and affective issues in information behavior. Data analysis entailed the use of the constant comparison method and descriptive statistics.

Data analysis led to the creation of a theoretical model of personal digital preservation practices. The model shows the effects of social, cognitive, and affective conditions on personal preservation decisions, as well as the effects of memory loss and technological advances over time, combined with information escalation over time. Because the preservation of personal digital information is the

result of personal, social, and technological interactions, the integration of these factors in the study of digital preservation practices is necessary for a viable solution to the digital preservation problem. The dissertation concludes with a discussion of implications for future research.

1. Introduction

Digital information is threatened by technology obsolescence. Users of digital information depend upon digital devices, computer hardware and/or software to view or hear the digital objects. Changes or advances in computer or digital technology can render digital objects or files unreadable or unusable. Libraries and other memory institutions are very aware of the crisis in digital preservation and are taking steps to preserve our collective cultural heritage (LeFurgy, 2005; Ross & Hedstrom, 2005). In contrast, Marshall et al's (2006) research suggests that individual consumers are much less aware of the impermanent state of their digital possessions, or if they are aware they feel disempowered to do anything about it. As a result, valuable representations of personal memories intended for future generations will be lost through ignorance and/or benign neglect (Yakel, 2004), and representations of family and social histories will be lost to what has been called the "digital dark ages" (Kuny, 1998).

Today many individuals amass large amounts of digital content because, like libraries, they have access to inexpensive and seemingly endless storage capability and to the high-powered computing needed to facilitate the creation and the downloading of digital content (Beagrie, 2005). However, the personal digital device environment offers limited support for content organization and preservation, so the likelihood that individuals will lose valuable representations of personal memories is very real (Jones, 2007).

As considered by this research, digital information is that information which is either born digital or has been digitized. An information item that is born digital is

one that was created initially in an electronic environment (Reitz, 2007). An information item that has been digitized is one that has been converted into a digital form from an analog form using a scanner or other conversion device (Reitz, 2007). Some items born digitally can be converted to a physical medium, for example, a Microsoft Word document can be printed out on paper. Other born-digital items cannot be wholly experienced outside of the digital environment, such as a website with its numerous links to other web pages and websites. Digitized information usually inherently has a hard copy backup, making it much less threatened by technological obsolescence than information that is born digital. Given that a recent study found that 93% of all new information is born digital and resides only on magnetic or optical storage devices (Lyman & Varian, 2003), the longevity of such information is the most threatened and in the most need of research and development.

This research project investigated information behavior related to the preservation of both born digital and digitized personal digital content. The following dimensions of preservation and personal digital information were examined: what steps are individuals taking towards digital preservation; affective responses to digital information; how preservation decisions are informed; and value and digital information. Additionally, this research explored individual experiences as they relate to digital content and physical information as representations of personal memories. The theory of information source horizons was used to explore the relationship between physical and digital information and to elicit the criteria participants use to determine the value of their personal information (Savolainen & Kari, 2004; Savolainen, 2007; Savolainen, 2008).

The focus of this research is the individual's relationship to personal digital information, particularly that content which the individual desires to preserve for the mid to long term, five years to 25 years and beyond. This time range was selected as it is difficult for individuals to anticipate specific future information needs (Bruce, 2004). Five years is long enough into the future to require - digital preservation for future accessibility, and 25 years represents a generation, implying the saving of digital items for the next generation. The timeframe encourages thinking about personal information beyond immediate uses and promotes thinking about the organization and maintenance of personal information over the long term.

The personal information studied in this research is the type of information that individuals keep for their own purposes, rather than the type of information kept by institutions about an individual. Examples of personal information kept by individuals include photographs, letters, emails, music, address book, websites, diaries, family trees, videos, and personally relevant news and journal articles, and may also include financial, legal, and medical documents.

While digital files and formats are vital to understanding digital preservation practice, personal digital information is discussed from the more holistic viewpoint of digital possessions, digital content, or personal digital information. Terms such as "digital object" and "digital artifact" will be avoided as these terms have specific meanings to the study of digital libraries and this research does not want to confuse technical terminology with more everyday expressions of digital information. In describing their experience of digital files, individuals are not likely to relate to their digital possessions in highly technical terms. Individuals want to preserve digital

information for the content and not necessarily for the formats, files and byte streams that comprise the information item. Of course, ultimately, individuals must achieve some level of understanding of the technical components of digital content for the sake preserving it.

Further, this research explored personal motivation for learning about digital preservation practices. The values associated with digital information (sentimental, emotional, financial, and historical) were explored as potential motivation. Specifically, the research considered the negative and positive cognitive and affective responses that individuals experience with personal digital information and what motivated them to learn more about organizing and preserving their personal information.

This research expands information behavior research beyond the information seeking process to include the study of information behaviors that contribute to the preservation of personal information. The research expands the study of affective issues in information behavior beyond the information seeking process through the study of affective responses to personal information possessions.

This research is important because it is situated in the context of the current transition in communication from the prominence of analog formats to digital. It is a time when the personal experiences of and relationships to information possessions and to representations of memories is changing from the tangible to the intangible, from those that are fixed in time and space to those that are mutable and fluid across space. Digital photos and digital music files represent two major cultural and technical shifts in the format of everyday information possessions. Until recently,

individual consumers and libraries alike (with some exceptions) have focused on the benefits of access and ease of content creation in the digital environment and have not fully considered the need to preserve digital content over the long term (Day, 1998; Alemneh et al, 2002). As most digital preservation research is focused outside of personal computing, personal digital devices, and personal collections, this research contributes to this area of digital preservation research and will benefit those individuals who wish to preserve personal digital information and collections for their own sake and the sake of their family history.

2. Review of Literature

2. 1 Theoretical Framework

The theoretical framework for this current research comes from information behavior research, a major area of interest within information science. Information behavior is defined in a variety of ways, but for the purposes of this research it is defined from the following four perspectives: *information behavior* in general terms is human behavior as it relates to information sources and channels regarding both active and passive information seeking and use; *information seeking behavior* is the intentional seeking of information to resolve an information need or to complete tasks or goals; *information searching behavior* is specifically the searching for information in information systems; and *information use behavior*s are the actions and processes that occur as information becomes part of an individual's knowledgebase (Wilson, 2000). By studying the behaviors that contribute to the preservation of personal digital information, this research intends to broaden this four-part definition to include the behaviors of personal information management and preservation.

Three major areas within the study of information behavior will be drawn upon: everyday life information seeking (ELIS), information horizons and information source horizons, and affective issues in information behavior. The framework is specifically based on the work of Savolainen, Kari, and Sonnenwald (Savolainen, 1995; Savolainen 2008); Sonnenwald, 1999; Savolainen & Kari, 2004). Savolainen's (1995) theory of everyday life information seeking (ELIS) was used along with Savolainen's and Kari's (2004) concept of information source horizons. Information source horizons is an extension of Sonnenwald's concept of information

horizons, which will also be used to frame this current research. In brief, these theories and concepts are chosen for three reasons. Firstly, the preservation and management of personal digital information, cognizant or not, is an everyday information activity that crosses the boundaries of work and non-work. Secondly, these theories and concepts can be extended to include information behaviors beyond information seeking. Lastly, the concept of information source horizons creates the opportunity to explore digital objects as unique entities and to understand the cognitive distinctions made between physical and digital personal objects.

2.1.1 Everyday Life Information Seeking (ELIS)

Bourdieu's (1984) theory of habitus informs Savolainen's ELIS theories (1995), Way of Life and Mastery of Life. Bourdieu views information seeking as a natural component of everyday practices; habitus is a socially and culturally determined system of thinking. One's everyday decisions are made --based on one's social class or cultural group. For this reason, this current research focuses on public library users. Individuals, who are primarily served by public libraries, rather than academic or corporate libraries, will have access to similar information services and are thus likely, but not certainly exist in a similar information environment.

Way of life is understood as the order of things. "Things" are all types of everyday activities related to work, household, and hobbies. "Order" refers to the preference given to these activities. Order is both objective and subjective; for example, objective order is time structured around work schedules and subjective order is represented by the activities voluntarily engaged in during leisure time. Disruption in the order of things is reflected in the distributed and ephemeral nature

of personal digital information. Digital information organization and access is fragmented through the use of computers accessed through work, home, and commercial web services. The networked nature of digital information creates an ostensibly fluid connection between home and work. Information fragmentation results, however, when an individual cannot recall where or how information was stored for later use (Jones, 2007).

Individuals possess a sense of cognitive order and know when things are normal. Everyday activities are typically managed, while behavior that is uncontrolled is exceptional. There is coherence to everyday activities giving individuals the ability to plan the meaningful use of their time. Three aspects taken as a whole characterize one's way of life: structure of time budget (work/leisure), models of consumption of goods and services, and nature of hobbies (Savolainen, 1995). These three aspects of way of life are ever-present in the use of personal computers.

Because the order of things is not guaranteed, individuals must actively care for the order. Mastery of life (MOL) is the caring activity and implies the importance of the life project at large. Understanding the steps necessary for digital preservation is a caring activity related to the development and documentation of the life project--. MOL can be active or passive. Passive MOL is when the order is progressing as expected and active MOL is when problem solving is needed to deal with threats to the order. "Mastery of life is a general preparedness to approach everyday problems in certain ways in accordance with one's values" (Savolainen, 1995, p.264). A positive MOL has a sense of coherence and predictability, a sense that one will know

how to deal with stimuli within one's environment to keep it stable and meaningful. Comprehensibility, manageability, and meaningfulness of stimuli are all part of the sense of coherence and therefore part of MOL. By studying the experiences that contribute to the preservation of personal digital information, this current research intends to explore - the sense of coherence - in the realm of tangible artifacts: is it unthinkingly being applied to digital possessions?

The cultural and social class one is born into, along with one's generation, creates the basic models of MOL. As most individuals who use public libraries (Vavrek, 2000; Japzon & Gong, 2005) and own personal computers are middle class, (Chakroborty & Bosman, 2005; Martin & Robinson, 2007) class will likely not be a factor in this study for contrasting - behaviors surrounding the preservation of personal digital information.

2.1.2 Information Horizons

Savolainen and Kari (2004) explore the significance of digital information in ELIS by examining to what extent individuals prioritize the internet as a research resource over other types of resources using the concept of information source horizons. Their research extends Sonnenwald's (1999) conceptualization of an information horizon. The concept is based on Sonnenwald's framework for understanding human information behavior. Sonnenwald's framework is based on theories and empirical studies in information science, communication, sociology, and psychology. Concepts of context, situation, and social networks and five propositions explain the framework, which describes information exploration, seeking, filtering, use, and communication. The framework incorporates social, cognitive, and systems

perspectives. Information horizon is a key concept. Within any situation and context, there is an information horizon in which an individual acts. The following illustration is an example from this current research: an individual who is functioning in the context of being a personal computer user and in the situation of deciding how to store personal information items over the long term consults information sources within her information horizon to make those decisions.

The information horizon framework is informed by information science research. Kuhlthau's (1991) uncertainty principle is included, as this views information behavior as a process influenced by both cognitive and affective responses. Wilson's (1997) general model of human information behavior is drawn from because it demonstrates the value of taking an inter-disciplinary approach. Belkin's anomalous state of knowledge (ASK) model (1980) is included for its contribution to understanding the behaviors surrounding the search for information in information retrieval (IR) systems. And lastly, Ingwersen's (1996) work is included because its perspective on situational and cognitive components are to be represented in IR systems for increased effectiveness.

Sonnenwald (1999) defines context as the embodiment of a collection of past, present, and future situations. Usually there is a shared understanding of a context by its participants. This shared understanding may not be identical or complete. Examples of context include: academic, family life, and citizenship. The boundaries of context are flexible. Contexts are multi-dimensional. Attributes of context include: place, time, goals, tasks, systems, situations, processes, organizations, and types of

participants. Characterizing context is difficult because contexts are not discrete entities – two or more contexts can share common attributes.

Situations and social networks are defined by Sonnenwald (1999) as follows. Within each context a flow of situations can arise. The context is larger than the situation. One context may exist for many situations. A situation is described as a set of related activities or set of related stories that occur over time. Individuals will describe the same situation differently. Individuals create patterns through communication with others which, create social networks. In turn, social networks help to construct situations and contexts and are likewise constructed by situations and contexts. Social networks exist within a context and situation. Social networks are not the focus of Sonnenwald's framework but are included owing to their influence on situations and context.

The following five propositions presented by Sonnenwald (1999) build on the concepts of context, situations, and social networks to build a framework of information behavior which focuses on the information horizon.

1. *Human Information behavior is woven around, i.e., is shaped by and shapes, individuals, social networks, situations, and contexts* (p.5). Given a situation and a context, an individual may encounter an information need and that same situation and context help determine the information need. Further, the individual, the situation, the context, and the social network may help determine the information resources available to satisfy that need.

2. *Individuals or systems within a particular situation and context, may perceive, reflect and/or evaluate change in others, self, and/or their environment.*

Information behavior is constructed amidst a flow of such reflections and/or evaluations, in particular amidst reflections and/or evaluations concerning a lack of knowledge (p. 6). There is always change and movement within a situation. The models of anomalous state of knowledge (ASK) (Belkin, 1980) and sense-making theory (Dervin, 1992) reflect this. The reflection on and evaluation of change creates the lack-of-knowledge condition. The reflection process is motivated by accommodating self, other, and the environment.

3. *Within a context and situation is an "information horizon" in which we act* (p. 8). When an individual seeks information, there is an information horizon in which they can seek information and it consists of a variety or resources. Information horizons are - determined socially and individually for situations and contexts.

4. *Human information behavior may, ideally, be viewed as collaboration among an individual and information resources* (p. 9). The goal of the collaboration is to share meaning and resolve the lack of knowledge condition. Reflexive interaction and/or reflexive provisioning of information are involved in the collaboration. Reflexive provisioning is when authors reflect on potential readers and then write. The reference interview is an example of reflexive interaction. The collaboration between an individual and information resources will be bounded by the individual's information horizon for that given situation and context.

5. *Information horizons may be conceptualized as densely- populated solution spaces* (p. 10). Individuals seek solutions to problems, needs, and goals from the diversity of information resources that are accessible to them.

The propositions for information horizons suggest that certain kinds of data are important for understanding human information behavior. Such data include: decisions being made; activities engaged in during information seeking; when and why a resource is used; including asking other people the relationships and connections between resources; individual preferences and evaluation of resources; and the impact of situations and contexts on the information seeking process (Sonnenwald & Wildemuth, 2001). These behaviors can be viewed in light of the processes individuals go through when they evaluate, organize, use/reuse, and preserve their own personal information.

Through the study of how individuals value personal digital information, this current research explores the information horizon that exists for personal information, in particular that information which individuals intend to access over the long term. The context is that of personal information as it relates to the whole of an individual's life. Rather than a situation that requires seeking resolution to an immediate or finite information need, the situation individuals are in, given the context, is one of predicating information needs and anticipating the value of personal digital information over the course of one's life.

2.1.3 Information Source Horizons

Savolainen (2006a) categorizes the information behavior research that uses spatial factors to explore contextual qualifiers in information seeking into three main approaches. In the *objectifying approach*, spatial factors are discrete entities that - constrain information seeking. In the *realistic-pragmatic approach*, objective constraints are acknowledged, but the reality of these constraints can be changed in

part and information behaviors altered. The *perspectivist approach* strongly emphasizes the subjective and situation-bound interpretation of spatial factors. In the perspectivist approach, the focus is on how people objectively assess the value of different sources by means of spatial constructs such as information horizons. The emphasis of the approach is on the meaning found in the construction of source preferences. For this reason, Savolainen and Kari (2004) selected Sonnenwald's concept of information horizons, from the perspectivist approach, for its emphasis on source preference.

However, Savolainen and Kari (2004) chose to adapt the concept to reflect a greater focus on an individual's information environment. Savolainen states that in Sonnenwald's concept of information horizons the relationship between information horizon and information resources remains ambiguous and synonymous (Savolainen, 2006a). They did not agree with Sonnenwald's view that books and libraries could be considered as information horizons in and of themselves. Savolainen and Kari (2004) developed the concept of information source horizons to include the differentiation of sources. The authors define information source horizons as an imaginary field which opens before the mind's eye of the information seeker. These horizons are created in the broader context of an information environment perceived of by the individual.

The horizon is created in the context of a larger information environment and not solely in the context of a particular situation as Sonnenwald has described it. When an individual is faced with solving a problem, she will construct an information source horizon from sources in her information environment. The process is selective and does not represent the entire information environment. Different criteria are used

to select and position sources on the horizon. Information source horizons are of two types: a stable horizon across situations indicating the ways in which people value information sources across situations and a dynamic problem or situation specific horizon. In terms of this current research, the stable information environment is the horizon of personal information, physical and digital, and the values associated with them. The situation specific horizon is the personal information objects that individuals want to preserve over the long term and the values associated with them.

While the theoretical underpinnings of both information horizons and information source horizons were used in support of this current research, -information source horizons was used in the methodological design rather than information horizons for two significant reasons. First, information source horizons represent an individual's personal knowledge and experience of the larger information environment. Second, the emphasis is on sources of information that can be identified specifically, rather than access points to information which can contain a variety of sources, as in a library. These two attributes of information source horizons support the goal of this research to understand what sources or objects of personal information individuals value.

2.1.4 Studies using Information Horizon and Information Source Horizons

In 2001, Sonnenwald and Wildemuth conducted a study in which participants were asked to describe and to draw their information horizons. Study participants included 11 undergraduate students (juniors and seniors ages 19 to 23) and nine corporate scientists involved in an electronic mentoring program. The students were

enrolled in a university that is historically minority and is located in a rural and economically depressed area.

The participants were asked to describe recent information seeking experiences within the particular context of science courses, science careers, and scientific work projects. The participants were asked follow-up questions to prompt for greater details of their experiences. Examples of follow-up questions include the following: what sources, what order and why, how the information was used, what would they do the same and what differently next time. Generally, participants were encouraged to talk about and explain their information horizon.

The data were analyzed by transferring the data to a matrix: the columns contained the students' names and the rows the sources used. The number in the cells of the matrix indicates the student's preference and/or order for each source. A network of the information resources for students was drawn and social network analysis performed. Connections and lack of connections among resources were illustrated.

The study participants were given a survey to identify the information resources they used in the past two months. The survey results indicated that the students used scholarly materials from printed indices and networked resources, books from the university catalog or from browsing the shelves. Interestingly, the information horizon maps showed that the students considered many more resources than information professionals traditionally consider in studies. Many of the sources used by the students are considered non-academic or leisure in nature: the Learning Channel, aunts and uncles, *Time* magazine, career centers, and hospitals. The authors

mention one student from another study that placed “God” several times on his information horizon, illustrating that it is almost impossible to consider all the sources. The information horizon technique captured more information about the process and relationship between information resources than did the survey questions. For the same reason, this current research combined interviews with a mapping technique.

Huliva’s (2009) research on the information seeking and use behavior of 25 Swedish and Finish archaeology professionals was informed by the theoretical framework of information horizons. At the start of each interview, participants drew a mental model of the information resources important to their daily work. The mental model was used in the manner of a checklist so that no resources were overlooked during the course of the interview. Huliva modified the participant-created information horizon maps to produce researcher-created analytical information horizon maps. In doing so, Huliva ordered and refined participant responses thereby making the responses comparable for analysis rather than rely on 25 informal and idiosyncratic responses for analysis. Analytical information horizon maps allow for a deeper analysis than do maps generated without theoretical perspective of the researcher.

Likewise, Savolainen’s and Kari’s (2004) study on how individuals prioritized the internet as an everyday information resource employed a mapping component. Their study focused on the ways in which information sources and channels, pathways on which information flows, are valued and prioritized in the context of ELIS. The study focused on the criteria that participants used to judge the quality of

sources and channels. Source accessibility and source quality are discussed as significant criteria of relevance. Eighteen (18) individuals participated in semi-structured interviews focusing on the role of the internet in self-development issues. Self-development is defined as the empowering of an individual's abilities, skills, etc. in the service or realizing one's personal potential.

The 18 participants were asked to place sources used for self development into zones. Zone one represents the most strongly preferred information sources; zone two represents those of secondary importance; and zone three represents peripheral information sources. They were given a diagram of three nested circles and were asked to treat the center circle as zone one. Also, the participants were asked to describe the reasons for their preference as they placed the items.

Quantitative methods were used to count types of information sources included in the horizons and qualitative methods were used to analyze the data received from the think aloud. The sources were grouped into six categories for analysis: networked sources, broadcast media, printed media, human sources, and organizational sources, and other sources. Of the 111 total sources mentioned, 26% were human sources, 23% printed media, 18 % networked sources, 12 % broadcast media, 10% organizational, and 12% as other. In zone one, human sources were the highest with 31% and next came networked sources with 29%. The reasons given for the use of the internet as a preference are ease of accessibility, currency, interactivity, and broad repertoire of information. In this current research the study participants expressed a preference for personal digital information for some of the same reasons. Interestingly zone one had the fewest types of sources, and zone three had the greatest

number of sources; whereas in this study, participants placed more information items in the center and fewer the second and third zones.

Points were given to each source according to zone. All sources in the first zone were given 3 points, second zone 2 points, and the third zone 1, reflecting the value of the sources. Using this scoring system, the internet or networked sources placed third after human sources (61) points, print media (55) and networked (46).

In two other studies, Savolainen (2007; 2008) used the information source horizon data collection technique described above. In the first study, he studied the information seeking preferences of 20 environmental activists specific to information sources and channels for - keeping current with daily news and events. In the second study, Savolainen analyzed the participants' information pathways in addition to their information source horizons. Information pathways indicate the path participants' take from one source to the next. He used these techniques to determine criteria for source and path preference when seeking information to solve specific problems.

In this current research, the three zones of the information source horizon were used to explore format preference and the value of personal information. Through the study of how individuals value personal digital information, this current research extended the information source horizons concept and mapping technique beyond the subdiscipline of information seeking to the study of preservation and value estimation within personal information management.

2.2 Affect Issues in Information Behavior

Important to Savolainen's concept of Mastery of Life is the way in which one deals with problem situations and seeks information to facilitate solving problems. Savolainen presents a typology for solving problems that is either cognitive or affective and optimistic or pessimistic, resulting in four types. In comparing cognitive and affective types, Savolainen presents a negative and common view of affective influence to problem solving. "A cognitive orientation emphasizes an analytic and systematic approach to problems whereas the affective orientation refers to the exact opposite: an emotionally laden and rather unpredictable reaction to the issues at hand (Savolainen, 1995, p. 265).

Much of information behavior research has focused on negative aspects of emotion in information seeking, in particular the anxiety surrounding technology and the research process (Mellon, 1986; Kuhlthau, 1991; Jiao & Onwuegbuzie, 1997); information overload (Wurman, 1989; Wurman, Sume & Loring, 2000); information overload at work (Allen & Wilson, 2003; Eppler & Mengis, 2004) and in everyday life (Jacoby, Speller, & Berning, 1974; Savolainen, 2006b). This line of research explores the extent to which positive affective responses, such as attachment, love, excitement, and joy for digital objects, affects behaviors in the digital information environment, and to what extent decisions regarding the preservation of digital objects are motivated by both negative and positive affective responses.

From current research in decision making, Isen (2004) discusses support for the view that affective response during decision making can have positive influence. When affect is involved in decision making, typically it is considered a disruptive

force which can lead to irrational and/or inappropriate decisions. Positive or mild feelings of affect are often not recognized as being a part of decision making because these feelings facilitate the process and are not noticed as a result. When other factors are equal, positive affect promotes exploration, enjoyment of new ideas, and new ways of seeing things. These influences are especially true in enjoyable or safe situations.

Support for the importance of emotions in individual pursuits comes from the field of neuroscience as well. Goleman (1995) asserts that the impact of emotions has not been fully understood as it relates to individual thought processes, decision making, and success. Based on neurological research, Goleman concludes that emotional intelligence can be the greatest predictor of individual success. Neurologist Damasio's (2006) research supports the importance of emotions and feelings in personal and social decision making. He specifically studies individuals who received damage to their frontal lobes, an area connected with emotional memories. The damage severed the connection to those memories leaving the individuals incapable of making decisions beneficial to mind, body and social self.

From -information behavior research, there is limited research and support for the positive influence of affective responses. James' and Nahl's (1996) research on adapting to Internet use found that affective responses, like confidence, excitement, attraction, and love, played a positive role in study participants' decisions to continue to use, learn, and adapt to the Internet. Their research supports the notion that affective responses can be both positive and influential. Nahl (1997) presents a counseling inventory to describe the connection between affective, cognitive, and

sensorimotor behaviors. She offers the inventory as means for instruction librarians to anticipate the dynamics between emotions, cognition, and actions when learning to use the internet. Awareness of what positive and negative affective responses correspond to which phases of learning is of value for communicating any process, from internet use to understanding and implementing digital preservation practice.

There is little research within the information science or information behavior fields that looks at the emotional connection to one particular information object. There is much in the way of studies that examine cognitive evaluations and connections to specific pieces of information, particularly user studies on relevance (Barry, 1994; Tang & Solomon, 1998; Barry & Schamber, 1998; Choi & Rasmussen, 2002; Xu & Chen, 2006). Buckland's (1991) discussion of information as thing explores the idea of information as an object apart from the idea of information as knowledge or as process. Buckland states, "if you can touch it or measure it directly, it is not knowledge, but must be some physical thing, possibly information-as-thing" (p.352). What happens to the concept of information as thing in the digital realm when information can no longer be touched or directly measured? Schamber (1996) questions the conceptualization of a document in view of electronic media. She contrasts electronic documents with physical documents and states in comparison that digital documents are easily manipulable, internally and externally linkable, readily transformable, inherently searchable, instantly transportable, and infinitely replicable. And goes on to state that "electronic documents seem to exist in the traditional sense only when individuals choose to transport or save information in some artifactual form" (p. 671). Through the study of how individuals value digital possessions, this

current research intends to explore the concept of document or information as thing as it relates to individual interest in the long-term preservation of personal digital possessions. Also, this current research explores the significance of emotional connections that individuals have for particular objects and any impact on the long term preservation it might have.

Further, information behavior research has focused primarily on affective issues related to information seeking and use. In *Information and Emotion* edited by Nahl and Bilal (2007), an emergent affective paradigm in information behavior research and theory is presented. Examples of research in support of this paradigm include: information seeking and the emotions of blind individuals, the social-emotional issues related to information literacy, library and research anxiety, and affective issues in information interactions in critical care nursing. There is no discussion or research included that addresses the affective issues related to personal information management, evaluation and/or appraisal. This current research has the potential to add to this paradigm as it provides support for cognitive and affective issues related to building and preserving personal information collections.

In contrast, in the book *Evocative Objects: The Things We Think With,* edited by (Turkle, ed., 2007), individual authors explore the emotional and cognitive connections to particular objects, physical, digital, and the representation of objects such as photographs. The tension between the world of physical and digital objects is particularly acute in the essay by Yee, who describes her full-sensory experience of working in the archives of the world-renowned architect Le Corbusier, and her disappointed reaction to the news that the archives will be digitized in its entirety,

thereby removing the need to interact physically with the archive. Yee comments on "how easy it was to trade the value of touch and physicality for the powers of digitization" (p. 34). She goes on to say that she "felt fortunate to be in a generation of designers that straddles both physical and digital worlds, a generation that creates, values and understands handmade drawings and models as well as digital ones" (p. 35). Through the study of the experiences that contribute to the preservation of personal digital information and how individuals value digital possessions, this current research intends to likewise explore the value of physical and digital objects, the relationship between the two and the implications for information preservation and value.

2.3 Personal Information Management

The 2007 *ARIST* chapter by Jones on personal information management (PIM) is the first review for this relatively new research area. Vannevar Bush is considered to have started the modern dialogue regarding PIM with his essay, "As We May Think"(1945). The idea of the memex machine brought to life the notion of having facile access to all of one's personal memories. PIM includes the management of information going into our memories, as well as external information. There are implications regarding how the human memory works for the successful recall and recognition of personal information items from organizational systems, computerized or otherwise (Lansdale, 1988).

Jones defines PIM as both the practice and the study of the "activities a person performs in order to acquire or create, store, organize, maintain, retrieve, use and distribute the information needed to complete tasks (work-related or not) and fulfill

various roles and responsibilities (for example as parent, employee, friend, or community member)" (p.453). Boardman and Sasse (2004) define PIM in the following terms:

> A prime characteristic of human behavior is to acquire and keep items of value. In both the physical and digital domains, our personal spaces become populated with the objects we accumulate as our lives unfold. Personal Information Management (PIM) is an umbrella term used to describe the collection, storage, organization and retrieval of digital objects (e.g. files, addresses, and bookmarks), by an individual in their personal computing environment" (p. 583). Bellotti and Smith (2000) define PIM as "the practice of managing the information that helps us in our daily lives (p. 227).

In support of this current research, the definition of PIM is expanded to include the preservation of personal information and to include affective issues in addition to the behavioral and cognitive. In so defining PIM, a bridge can be made between information behavior research and PIM research. Information behavior research is a principal domain of library and information science research, and from this view much of what is considered in - PIM has been thought of in terms of the work of the formal institutions. Libraries and other information providers do the work of collecting, organizing, maintaining, storing etc. (Svenonius, 2000). The focus of information behavior research has been on the seeking and searching and the using of information from systems that organize, maintain, and store information. Whereas it appears that within the field of Human Computer Interaction, researchers have been

more attuned to the human behaviors related to personal computing, as evidenced by the PIM conference has taken place via the CHI workshop venue since 2005.

Furthermore, personal information management is studied from three different perspectives. First, information kept by individuals for their own personal uses. Secondly, information kept about individuals by others, for example, legal, medical, and financial information. Thirdly, information experienced by individuals involuntarily, for example, viewing web pages and browsing print material (Jones, 2007). This current research will focus on the first perspective of PIM, information kept by individuals for their own personal uses.

PIM research emphasizes organization and maintenance issues. While there are tools for helping individuals organize personal information, the problem of information fragmentation persists and can be exacerbated by the diversity of tools that exist. Information fragmentation is the result of using different organizational schemes for different types of information; for example, organizing documents, emails, and photographs in different ways and in different places. Information fragmentation is compounded by the scattering of information into information islands, digital information supported/accessed through diverse applications and devices (Jones, 2007).

Several studies have focused on the effectiveness and potential of PIM tools. The four major areas under study are: web site management tools (Abrams, Baecker, &Chignell, 1998; Gottlieb & Dilevko, 2001 Dix & Marshall, 2003; Bruce, Jones, & Dumais, 2004); files and folders (Carroll, 1982; Barreau, 1995; Barreau & Nardi, 1995; Jones et al 2005); email (Whittaker & Sidner, 1996; Balter, 2000; Ducheneaut

& Bellotti, 2001; Marshall, 2006; Whittaker, Bellotti, & Gwizdka, 2006),;photographs (Rodden & Wood 2003; Cunningham & Masoodian, 2007); and cross-tool studies (Boardman & Sasse, 2004; Taveen, 2004; Karger & Jones, 2006). The collective goal of the aforementioned research is to understand how these tools are used by individuals to manage personal information and to make recommendations for system and/or tool improvements based on those observations.

Jones (2007) asserts that advances in PIM tools will allow individuals to make better decisions regarding their resources, time, energy, and attention, and thereby improving their quality of life. Advances in PIM research will result in new and/or improved tools as well as new teachable techniques of information management via information literacy programs. In Chapter Seven, implications of this current research for teachable methods of PIM are discussed.

The concepts of personal information space (PSI) and personal information collection (PIC) are key PIM concepts to this current research. A PSI includes all the information that is under an individual's control: books, documents, electronic bookmarks, email messages, etc. One can only have one PSI (Jones, 2007). This concept is similar to information source horizons with the exception that information source horizons include information not under an individual's control, such as other individuals. PICs are islands in the PSI. PICs result from the conscious effort to control the information that goes into the collection and the manner in which it is organized (Jones, 2007). A PIC is a subset of an individual's information world, -used when an information need occurs (Bruce, 2005). In this way, PSIs and PICs relate to the two types of information source horizons: PSI is like the stable horizon

that exists across situations indicating the ways in which people value information generally; a dynamic problem or situation specific horizon is likened to the PIC. These are relationships are not exact but do complement one another.

A framework for studying PIM can be taken from Barreau's (1995) research on the PIM behaviors in the workplace that involve the use of a personal computer. She studied the PIM behaviors of seven managers through observation and interviews. Barreau asked that they give a guided tour of their electronic directories to identify documents in their workspace and to identify reasons affecting their decision to keep the information in their workspace. She compared electronic behaviors to the behaviors Kwasnik (1989) observed in the behavior surrounding physical documents. She concludes that hardware and software may influence behaviors, wwhich in turn may impact personal and organization efficiency. The increased flexibility of software over print document organization tools creates the opportunity for personalized and unique information behaviors. Additionally, she found that the physical cues of print document storage have been replaced by the personal and situational cues in the digital realm. Cues are important to the recall of stored information.

Barreau framed her research in terms of four overarching behavioral areas of PIM: acquisition, organization, re-finding and/or retrieval, and maintenance. The acquisition of an item can be deliberate or unintentional, for example, saving a file versus receiving an email. Acquisitions that are deliberate reflect a personal anticipated information need (Bruce, 2005). Individuals acquire an item with the intent to use the item in the future. In Jones' (2004) research on re-finding behaviors,

he found that individuals will either acquire (keep) the information into a PIC, leave the information where it is for finding again, or ignore the information altogether.

Keeping or acquiring information items requires more organizational effort upfront if individuals want to save time later when they attempt to re-find/retrieve information they know is in their PSI. If an individual remembers that an information item is in their PSI, then it likely because of some explicit act of keeping (Jones, 2004). "The two steps of information recall and recognition can be viewed as a dialogue between people and their information environments." (Lansdale, 1988, p. 468) The success of finding/retrieving information in a PSI depends on the individual remembering what to look for and where to look for it. Remembering that an item is owned must precede recall.

Relevant to the next PIM behavior, organization, is the digital memories and the record everything movement (Cutrell, Dumais, & Teevan, 2006; Gemmell, Bell, & Lueder, 2006). Such projects as Stuff I've Seen and MyLifeBits illustrate that "everything" can be recorded and stored in a continual and comprehensive manner. The creators of these projects advocate saving everything and relying on search for re-finding personal information. Providing for information organization through metadata and file folders is considered unnecessary, given the powers of search (Cutrell, Dumais, & Teevan, 2006).

In addition to the cognitive associations built around acquiring a personal information item, memory is aided by keeping items in view. A study by Barreau and Nardi (1995) found that users of personal computers purposely locate items to facilitate remembering, for example, using icons on desktops and sending email

messages to oneself. Gemmell, Bell, and Lueder (2006), advocates of search over organization, provide the screen saver as the "killer app" to manage the recall and recognition problems that will result from having an overabundance of personal digital objects. Digital representations of an individual's life will randomly pop up on her screen saver to remind her of what is stored on her personal computer.

As items increase and PIM tools diversify, it becomes more difficult to manage or see all items. Over-reliance on search, in place of attempting to organize information, may not be that helpful in re-finding information over the long term. Jones recommends creating a personal unifying taxonomy to integrate information organization across devices (Jones, 2004). Marshall (2008) recommends using a federated search mechanism to unify access to metadata applied to personal digital information rather than creating a single repository of all personal digital content. Anderson, Hodge, and Japzon (2007) provide a framework for a metadata repository that combines the application of a unifying taxonomy and a federated search to the distributed networked content of a large research organization.

Also, this study and others have shown a preference for browsing over searching (Barreau, 1995; Boardman & Sasse, 2004). The involvement of the individual in the act of organizing her information invokes the use of memory that will then later aid the recall and recognition of information over time (Lansdale, 1988). Even if a search tool recalls something, the individual may not recognize relevant information without previously established cognitive associations to it (Lansdale, 1988). As mentioned previously, Barreau (1995) found, in the digital realm, that personal and situational cues aid retrieval of items from personal

collections. As Marshall (2008) points out, tools are needed to promote the "re-encountering" of personal information that is buried in one's decade deep personal stores of information

It's important, though, to discuss the possibility of saving everything because the converse of saving of everything without a thought given to organization is the saving items with care. If individuals were to consider the value of unique digital items, they would incorporate a curatorial process to the digital collections they are creating. The curatorial process would encourage individuals to evaluate and to make judgments on what items are of particular importance and which deserve more care over the long term. In a study by Petrelli, van den Hoven, and Whittacker (2009) ten families were asked to create time capsules of personally valuable information that would cue their memories in the distant future. These families did not want an exhaustive record of their digital lives but rather preferred to include carefully selected cues that were more often in a physical format.

This current research explores how individuals think about digital items as unique possessions. PIM studies focus on the collections of information items, whereas this current research explores the emotions attached to and the behaviors that surround particular individual objects in the digital realm. Through the study of how individuals value digital possessions, this current research investigates what individuals consider and reflect on regarding curatorial aspects of building collections.

Re-finding and/or the act of retrieving information from personal collections is another behavior studied in PIM. Much of information behavior research focuses

on the processes involved in finding information in the first place. So PIM research takes up where information behavior research leaves off. However, the Krikelas' (1983) model of information seeking does acknowledge that most individuals seek information internally, from their personal stores of information. Given the context of this model, a PIC is a likely information source for individuals to find or retrieve information to complete a task or solve a problem (Bruce, 2005), thereby creating a connection between the two fields of study.

PIM studies have found that maintenance and organization of information is less of a priority to individuals than time-sensitive and context-driven activities such as finding and keeping (Barreau, 1995; Bruce & Jones, 2004). Maintenance activities include storing, deleting, and reorganizing information. In the PIM literature, there is limited research on the behaviors that surround the maintenance issue of digital preservation except for the work by Marshall (2006, 2008), Marshall, et al (2006), and Petrelli, van den Hoven, and Whittaker (2009). It should be noted that in the archival literature, the self-archiving of electronic personal records by scholars, artists, academics, and politicians is being researched (Kaye, 2006; Kim, 2007), as is the ingestion of personal electronic records into existing institutional archives (Cunningham, 1999; Williams, 2008; Davis, 2008).

Marshall et al (2006) explored individual attitudes towards and awareness of digital preservation. Their informal study involved interviewing twelve participants to determine what kinds of digital possessions the participants kept, what they cared about over the long term, and what barriers they encountered to preservation. The study participants were sophisticated digital device users and had amassed a diversity

of digital object types and an abundance of digital personal objects. Most of the participants had experienced the loss of valuable digital objects.

Their research exposed gaps, contradictions, and falsities in the study participants' knowledge regarding digital preservation and, from this, the researchers identified five folk principles: replicating, culling, keeping, losing, and replacing. Participants mistakenly thought of replicating or backing up digital files as a sufficient means of long-term preservation. Owing to the inexpensive and abundant nature of storage, participants were reluctant to delete anything or to differentiate between active memory and long term storage. Participants prefer to cull everything together in one large mass. Keeping refers to the value-neutral stance that the participants displayed regarding the items they kept. This value-neutral stance was apparent in the large numbers of email that many of the participants kept without a concern for backing them up. The participants accepted the eventuality of losing digital as part of the transient nature of digital information. Digital objects that were not created by the participants, but were acquired from other sources, were considered replaceable and not as important as personal pictures and documents.

Marshall, et al (2006) offer the following measurements of value for digital assets. A digital object's worth can be assessed in the following ways: the more times an object is replicated the more valued it is; the type of object it is and the mode of creation can influence degree of value; the amount of time spent creating an object can increase the value; the cost and stability of an object can determine value; and the emotional impact that an item can possess, as represented by it having been shared.

This current research likewise examined the value attributed to digital content and the attitudes towards and awareness of digital preservation, and also examined some additional aspects of personal computing and digital devices and digital preservation. This current research provides a unique contribution to both - personal information management and information behavior research by providing insights into understanding the following: motivations for learning about digital preservation practice; the specific emotions and behaviors that are involved with the preservation of one particular digital item; awareness of the role of digital formats in personal information preservation; and the relationship between physical and digital items in personal information preservation and information preservation generally.

Marshall (2006) conducted a case study using email correspondence between two individuals to explore the issues inherit in the long term storage, preservation, and access of personal information archives. She states that while technical issues of digital preservation are the same for the individual and the institution, personal information archives present unique problems. Many of these issues were previously mentioned in this section: such as digital content amassing quickly, difficulty in anticipating value of objects, the seemingly endless boundaries of a digital object, the time it requires to organize and curate materials, the specialized skills and knowledge necessary to maintain a digital collections, and issues surrounding privacy and networked access. There are implications of this study for teaching public library users the technical aspects of maintaining and preserving their own personal digital collections.

Marshall (2008b) presents an overview of the challenges faced by individuals based on five different sets of informal interviews along with her personal insights into the problem of personal digital preservation. She perceives the four biggest challenges facing individuals to be digital stewardship, distributed assets, value and accumulation, and retrieval from long term storage. This current research addresses each of these points

2.4 Digital Preservation Practice

Discussions of the digital preservation crisis and the loss of digital information are taking place in a variety of venues, including newspapers, scholarly and trade journals, Microsoft technical reports, and NARA and Library of Congress publications and initiatives. *The New York Times,* the *Washington Post* and other popular news publications have run articles on the topic (Hafner, 2007; Barksdale & Berman, 2007). Kuny in 1998 was the first to use the term "Digital Dark Ages" to liken today's loss of the digital written record to the Middle Ages, a period for which little written records have survived. Kuny's address to the IFLA council articulates the state of digital preservation and the problems that existed then and which continue to grow now. Hedstrom (1997) calls "digital preservation a time bomb for digital libraries" (p. 189) and the Council on Library and Information Resources referred to digital preservation as technological quicksand (Rothenberg, 1998).

Since Kuny's address steps have been taken by the library and information science community to better understand digital preservation. Examples include the beginning of the development of a format registry under the auspice of the Digital

Library Federation in 2002 and the PREMIS Data Dictionary, a joint effort of OCLC and RLG which began in 2000.[1] [2]These two examples address the importance of describing and recording the technical aspects or content of digital information to digital preservation.

This section will explore the landscape that surrounds digital preservation and technical representation of digital objects and the challenges that are being addressed by the digital library community. Of interest to this current research is that these same challenges exist for individuals and the preservation of their personal digital information. The preservation issues for institutional digital libraries are of a magnitude far greater than the personal digital collections of individuals, but the resolution for both for will ultimately be the same. To a certain degree, individuals must come to understand the instability of digital content and to acquire and use knowledge of digital preservation practices in order to preserve their own collections.

The purpose of this section is two fold. First, it is to establish connections between the steps taken by the library community towards improving the state of digital preservation and the ones individuals need to consider while creating personal digital collections. Secondly, it is to illustrate the complexity that surrounds digital preservation and the implications related to the average individual's understanding of digital preservation practices and for the design of personal computers and other digital devices and web services that support collections of digital content.

The technical representation of digital content is arguably the most critical aspect of digital preservation. However, before focusing on this aspect, the

[1] http://hul.harvard.edu/gdfr/
[2] http://www.oclc.org/research/projects/pmwg/premis-dd.pdf

dissertation will briefly discuss the importance of digital storage media and how it relates to and is often confused with the act of describing and documenting the technical information that is integral to the rendering of digital content. Digital media is fragile and given to corruption (Ross & Hedstrom, 2005) and when compared to acid-free paper which can survive for several hundred years if kept in the right conditions, most digital media falls short, lasting 5 years or less (Borgman, 2003a). For this reason, it is recommended that individuals and organizations create redundant copies of digital files on multiple storage devices (Hedstrom, 1997). Planning for redundancy creates a cushion around digital content. A hard drive could crash, a compact disc or other portable storage device could become corrupted, and with redundant files copied to an external hard drive or on networked server space, files identical to those lost can easily be retrieved.

Creating multiple copies of digital files in preparation for those digital files to be damaged or corrupted is a short- to mid-term preservation solution and it should not be confused with a long-term solution. As Chen (2001) suggests, the paradox of digital preservation is that if one wants to continue to have access to digital content, s/he must not hold on to it. If one does not make modifications to digital content, over time access to it will be lost. Chen states, "on the one hand, we want to maintain digital information intact as it was created; on the other we want to access this information dynamically and with the most advanced tools" (p. 25).

The LOCKSS, Lots of Copies Keep Stuff Save, organization makes an attempt to address the unique position libraries are in regarding the preservation of electronic serial subscriptions which are not owned but licensed by libraries. At the

same time, publishers providing journal subscriptions are not in the business of providing access to journals in perpetuity.[3] The LOCKSS example illustrates one of the ways in which legal and social issues also contribute to the digital preservation crisis. Whose responsibility is it to preserve digital content and to pay for the ongoing maintenance required for the preservation of digital content? As yet, it is not possible for digital preservation to be a public good as research dollars and marketplace rewards for research investments are needed to deliver digital preservation solutions (Workshop on Research Challenges in Digital Archiving and Long-term Preservation, 2004).

The individual consumer is faced with issues similar to digital serial subscriptions when dealing with online web services, such as the popular digital photograph storage sites, Flickr and Snapfish, and free email services, Hotmail and Yahoo. Is it the responsibility of these vendors to provide perpetual access to digital photographs and digital correspondence? Is the same service that provides access and organization today expected to provide access 10, 15 or 100 years from now? Those advancing the state of the art in digital preservation assert that in the digital environment access, organization, and long term preservation must be part of the same system/service design for digital preservation to ever be a reality.

Ideally, public library users and other consumers would have easily accessible and understandable information available to them on the most reliable and most appropriate storage devices and locations. A storage selection guide is part of a digital preservation workbook created for professional archivists by the United Kingdom Research project, PARADIGM (2008). Also, agencies such as National Institute of

[3] http://www.lockss.org/lockss/Home

Standards and Technology, Image Permanence Institute, and InterPARES have produces guides on the care and handling of digital storage media for professional audiences. While a variety of research has been conducted on the longevity of storage media (Porck & Teygeler, 2000) and a variety of guides produced, a consumer level information guide describing and comparing storage media does not exist.

2.4.1 OAIS Model

The Open Archival Information System (OAIS) model was designed by the Consultative Committee for Space Data Systems as part of an International Organization for Standardization (ISO) initiative to develop standards for the preservation of satellite data. Extending beyond its original purpose, the model has become a *de facto* standard in digital preservation by offering concepts and terminology to be shared across information cultures, thereby creating common ground for collaboration (OCLC/RLG Working Group on Preservation Metadata, 2002). In the OAIS model, what is referred to as *Representation Information* is discussed as preservation metadata in the literature. Representation Information makes the accurate rendering, understanding, interpretation of a digital object's content possible.

The OAIS Model illustrates a system design that addresses the simultaneous need for access and preservation. Long-term access requires the cooperation of social and technical systems. Any preservation metadata purposefully recorded or automatically extracted would have to function within an archival system. This is especially problematic for the individual given that the designs of the systems that

drive personal computing and digital devices are not archival in nature. The OAIS Reference Model is a high-level conceptual framework that describes the features necessary for an archival system to function. The OAIS model does not prescribe system requirements or implementation procedures. It is a theoretical model that describes the broad types of information or metadata required to both preserve and access digital content stored in an archival repository (Day, 2004).

2.4.2 Preservation Metadata

Preservation metadata has been defined by various digital preservation researchers. Several of these definitions will be provided for consideration. Besser (2000) indicates that preservation metadata is a means for providing the technical information necessary to support the two primary means of digital preservation: migration and emulation. Preservation metadata should capture and explain the technical environment needed to view particular digital content. This includes the specific hardware and software versions, decompression schemes, and other related files needed to view the digital content.

The OCLC/RLG Working Group on Preservation Metadata (2002) states that, "preservation metadata is the information infrastructure that supports the processes associated with digital preservation. More specifically, it is the information necessary to maintain the viability, renderability and understandability of digital resources over the long-term" (p.1). Viability requires that the digital object's bit stream is intact and readable from the digital medium upon which it is stored. Renderability is the translation of bit streams into a form that can be viewed and used by humans and

computers alike. Understandability requires that enough information be provided with the rendered object such that it can be understood by human users.

The *Best Practice Guidelines for Digital Collections* authored by the Office of Digital Collections and Research at the University of Maryland (2007) offers this definition of preservation metadata: "supports long-term retention of the digital object and may overlap with technical, administrative, and structural metadata." The only metadata schema omitted is the descriptive metadata which deals with the intellectual content of a digital object. As indicated previously, over time intellectual access is just as vital as technical access- for finding and reusing digital objects (Jones, 2004). The other metadata types presented in the definition of preservation metadata are as follows: administrative, which facilitates management of digital and analog resources; technical, which describes the technical aspects of the digital object; and structural, which describes the relationships within a digital object.

Day (1998), from the UK Office for Library and Information Networking, adds to the *Best Practices Guidelines*: metadata reflecting the rights to preserve and access digital information objects as custodial organizations often do not have physical custody over digital objects created by others, such as authors and publishers. This is important for individuals as well as custodial organizations. Individuals incorporate journal articles, images, music and other artifacts into their own personal digital collections, creating a new context for the information (Borgman, 2003b) and at times leaving provenance behind.

Day extends Besser's (2004) two primary means of digital preservation to four different strategies: preserving technology, emulation, migration, and

encapsulation. Preserving technology does not require generating metadata but keeping alive, as it were, obsolete hardware, operating systems, and software. For obvious reasons, this is not seen as a practical solution by many. The three other strategies require the capture, creation and maintenance of metadata. Preservation metadata is what will allow for the possible re-creation of digital objects as newer technologies replace obsolescent technologies.

2.4.3 Emulation, Migration, and Encapsulation

The next section will briefly describe the concepts of emulation, migration, and encapsulation, and the implications for preservation metadata and long-term access to personal digital objects. Emulation involves the development of programs that mimic the functions of obsolete programs on new platforms. Preservation metadata associated with objects from obsolete programs provide the technical context needed to run the programs and render the objects viewable or usable. Migration involves the transfer of digital objects from one generation of computer technology to the next. This is the most common preservation strategy employed. Encapsulation is based on the notion that preserved objects should be self-describing. The digital content is encapsulated with all the information required for the content to be deciphered and understood. Encapsulation is one of the basic concepts of the OAIS model and can be used to support either migration or emulation (Day, 2004). In the OAIS model, the Information Object is the encapsulation of the Data Object and the Representation Information. The Data Object is the actual content under consideration for access and preservation and the Representation Information is what

imparts meaning to the object's bit stream (OCLC/RLG Working Group on Preservation Metadata, 2002).

Day and Alemneh et al (1998, 1997, 2002) assert that preservation metadata has been an after-thought in creation of digital collections when, in truth, the capture of preservation metadata needs to begin with the creation of each digital object. The technical information needed to support each digital object should be part of its creation because, without that technical infrastructure, the object will not remain accessible over the long term. Encapsulation provides the means to integrate the content with its technical infrastructure. Initially, digital collection creators gave most thought to describing content to promote content discovery. As digital collections and digital content acquisitions continue to grow in size, institutions are beginning to plan and budget for the preservation of the digital collections and content they are acquiring (Searle & Thompson, 2003). Institutions are beginning to realize that digital content cannot be acquired without the simultaneous consideration of preserving that content. Through the study of how individuals are motivated to acquire knowledge for digital preservation, this current research intends to explore the extent to which individuals are motivated to learn about simultaneously creating/acquiring and preserving their own digital information.

As previously stated, the research by Marshall (2006, 2008) and Marshall et al (2006) is the only research that has begun to explore digital preservation in the personal computing environment. However, much research related to digital preservation is being conducted within the digital libraries community. Ross and Hedstrom (2005) point out that while much imaginative and groundbreaking research

has been performed in the name of digital preservation, no path has been cleared to meet the challenges of digital preservation. Though, they do admit that current research efforts do have the promise of potential. The following sections describe aspects of the current research: PREMIS, the Global Digital Format Registry, METS, preferred formats.

2.4.4 PREMIS

In 2003, OCLC and RLG created an international working group to develop a core set of metadata elements for digital preservation. The Preservation Metadata: Implementation Strategies (PREMIS) working group created a set of metadata elements that are implementation independent, practically oriented, and likely to be used in most repositories. The working group sought to establish best practices for implementing preservation metadata and to expand the conceptual structure OAIS by mapping the metadata elements to its structure (Caplan & Guenther, 2005).

In 2005, the PREMIS working group released the data dictionary, their final report, and examples of data dictionary use.[4] The core elements are designed to ensure the preservation of object viability, renderability, understandability, authenticity, and identity. The group decided on the term "semantic unit" to represent metadata elements in the data dictionary. Semantic units represent information concepts generally rather than those specific to any digital object. Semantic units were compiled based on specifications proposed for the Global Digital Format Registry.

[4] http://www.oclc.org/research/projects/pmwg/

The group created a data model to give a sense of how the semantic units might be applied to specific objects. Also, the data model created a structure for the data dictionary. Five types of entities created the data model: event, agent, right, intellectual, and object. Semantic units were created for all entities except the intellectual and agent entities as these entities are not directly related to digital preservation. The object entity was divided into four subtypes: file, filestream, bitstream, and representation. The subtypes defined by the working group describe the technical infrastructure of an object. File is the most commonly understood of the subtypes. A file is accessible by a software application through an operating system. All files have a format which is defined by an established structure that specifies how the data are organized. A file may contain zero or more bitstreams and zero or more filestreams.

A bitstream is data within a file that cannot be transformed as a standalone file without additional structure being added to the stream. A filestream is a continuous set of bits that can be transformed into a standalone file without additional structure being added or reformatting of the stream. These distinctions are useful for understanding how best to digitally preserve all the components of complex or multi-file representation of intellectual entities. For example, a bitstream is an image embedded within a PDF and a filestream is a TIFF image within TAR file (Caplan & Guenther, 2005). The PREMIS working groups provides several examples that illustrate the use of semantic units described in the data dictionary to describe real digital objects.[5]

[5] http://www.oclc.org/research/projects/pmwg/premis-examples.pdf

2.4.5 Global Digital Format Registry

In 2005, the Global Digital Format Registry (GDFR), which began as a Digital Library Federation investigation in 2002, became a Harvard University Libraries' two-year project funded by a $600,000 grant from the Andrew W. Mellon Foundation. The purpose of the GDFR is to maintain a distributed service that provides storage, discovery, and - delivery representation information about digital formats. The registry will be a repository for format representation information, which is the descriptive, administrative, and technical metadata about digital formats. The registry will also include syntactic and semantic characteristics of the formats. It should be noted that the National Archives of the United Kingdom is also creating a global format directory, PRONOM. [6]

With this type of information, the GDFR should be able to perform the functions necessary to answer the types of questions common to the maintenance of a digital preservation repository (Abrams, 2005). The functions include: identification, validation, characterization, processing, and risk assessment. With these functions operational, individuals can identify unknown format types, determine the properties of a format, and check for technological obsolescence.

The creators of the GDFR and of the PREMIS data dictionary are aligned in the terminology that they use to describe digital formats. In fact, global "digital format" registry was selected instead of global "digital file" registry to capture all levels of granularity of content streams. The GDFR will describe details of specific format types and the PREMIS data dictionary will provide the means to consistently

[6] http://www.nationalarchives.gov.uk/PRONOM/Default.aspx

and therefore meaningfully describe the semantic units of the digital formats contained within the registry. The work of both of these initiatives conceptually relates to the OAIS model. In this way, the technology, terminology and the implementation of both can align and be used across information environments.

2.4.6 METS

The Metadata Encoding and Transmission Standard (METS) schema is an initiative of the Digital Library Federation which is maintained by the Network Development and MARC Standards Office of the Library of Congress. METS takes the form of an XML document that encodes all the metadata necessary to manage digital objects within and between repositories. A METS document consists of five major sections: descriptive metadata, administrative metadata, file groups, structural map, and executable behaviors. A METS document is intended to provide access to the technical information needed to migrate and refresh the data associated with digital objects.[7]

A METS document can be thought of as a mechanism for contextualizing preservation metadata. The complex links between the different types of metadata for a particular digital object are expressed and maintained through the METS format. Further, the METS format can be used in the OAIS model to create and deliver the information packages that support both archiving and access in that model.

Discussion of the PREMIS data dictionary, the GDFR, and METS illustrates the level of technical knowledge and skill required for digital preservation practice. Of interest to this current research is the implication of this discussion for making

[7] http://www.loc.gov/standards/mets/METSOverview.v2.html

digital preservation practice intellectually accessible to a wide range of individuals, including information professionals who create digital collections. Of further interest, the discussion has implications for systems design of personal computing. Personal computers and other digital devices are not designed to support both archiving and access. This current research explored the behaviors and affective responses that exist given the absence of a systems infrastructure that supports long term and stable access to digital objects.

2.4.7 Preferred Formats

The aforementioned *Best Practice Guidelines for Digital Collections* from the University of Maryland Libraries differentiates between archival formats and web deliverable formats. Section 7.1 of the guide states the following general requirement for archival file formats: "a format that is open (non-proprietary) and well-documented, widely supported, and cross-platform compatible." The Electronic Literature Organization provides a pamphlet both online and in print that encourages authors to publish and work in Acid-Free Bits.[8] Their description of why open systems support long term access of digital information is provided here:

> An open system is one whose essential workings are fully, publicly documented; an open standard is published and available to anyone. Those who use open systems and adhere to open standards when creating electronic literature have a much better chance that the format of their literary works will be supported, or decipherable, in the future. The small group of people in charge of a closed system or standard may lose interest and stop developing software, or the small group may change the system or standard without warning, so that older works of electronic literature no longer work on new platforms. (This is particularly a risk when electronic literature is not the main purpose of a system, and may be obliterated incidentally.) Open systems and formats can be most

[8] http://eliterature.org/pad/afb.html

> easily migrated and emulated, since their specifications are publicly known. Closed systems are far more difficult to migrate and emulate.
>
> A closed system may provide important capabilities that are otherwise not available, and some closed systems may be very well suited for the type of literary creation in which authors are interested, so there may be good reasons for authors to use a particular closed system. However, authors should be aware that such a choice could affect the longevity of their works. As a result, authors may wish to document such projects more thoroughly.

In addition, the use of open systems and formats will allow for the creation of standards. In turn, standards will allow for and promote interoperability among systems which will ultimately increase the possibilities for resource discovery and the preservation of the resources. Recommendations on file formats best for long term access have been made by agencies, like the Digital Preservation Coalition in the United Kingdom and the Florida Center for Library Automation Digital Archive in the United States, which are developing digital archives based on open systems, interoperability, and lossless formats (files that have not been compressed to save space). Information on file formats preferred for long term access needs to be made accessible to consumers as well.

The preservation of digital information will ultimately mean the loss of information. Through the process of migration or transferring content from older to newer formats (proprietary or open source), the integrity of the digital content will be changed or lost to the technology process trying to preserve the content. It has been suggested that research in this area be conducted to understand fully the extent of loss involved with a variety of formats and how much loss is acceptable to the users and to overall integrity of the object under consideration (Workshop on Research Challenges

in Digital Archiving and Long-term Preservation, 2004). This concept of inherent loss calls into question the feasibility of maintaining object authenticity over time. Information documenting the provenance and fixity of any object becomes increasingly important where information as evidence or the need for original expression is concerned.

2.4.8 Digital Preservation and Personal Digital Devices

This literature review has discussed various aspects of digital preservation practice. It is by no means an exhaustive description of all that is involved. It provides a pathway to the current study by illustrating the complexity of knowledge involved with the practice of digital preservation and the need to find ways to make that knowledge more accessible so that individuals intending to keep personal digital possession over the long term can be empowered to act. It also illustrates the need to develop systems that can simultaneously archive and provide access to digital information objects in all computing environments and for all layers of digital information. Clearly, collaboration is taking place and needs to continue to accomplish the daunting task of providing for the future of digital information. Representing the technical content of digital information is ultimately as important as representing the subject or descriptive content of digital information. This needs to be done for every type of digital information: government, research, and personal. Digital preservation is an issue for all who have come to depend on digital information for communication practices and, therefore, knowledge and use of preservation practices needs to be pervasive.

Digital preservation is technically complex, expensive, and evolving. All these factors create barriers to individuals preserving digital representations of personal memories. This current research provides insights into the factors that motivate individuals to learn more about digital preservation and overcome these barriers.

Research Questions

The following three questions guide this research study of digital preservation practices of the public library user participants. These research questions are designed to explore the personal, social, and technical aspects of digital preservation.

1. What kinds of personal digital information do public library users collect and why?

2. What are the cognitive, affective, and psycho-social influences that contribute to the preservation of personal digital information?

3. How can individuals improve their digital preservation practices and what would motivate them to make improvements to these practices?

3. Methods

The goal of this research is to understand the behaviors, beliefs, experiences, and affective responses of individuals regarding personal information and digital preservation and to understand the implications of these factors for information value and the motivation to learn digital preservation practices. Qualitative research methods were utilized since such methods are best used when studying the beliefs, behaviors or cultural aspects of human phenomena and when the human behaviors or emotions under study are not fully measurable (Krathwohl, 1997).

In this research, multiple data collection methods were employed to obtain a variety of views on the behaviors and attitudes that contribute to digital preservation. The use of multiple methods allows for deep insights into the problem under study as it allows for the diverse realities that exist for each participant to be expressed and used to construct, strengthen and validate theory (Golafshani, 2003). Triangulation is achieved through the use of multiple methods in qualitative research. The quantitative research methods measures of reliability and validity do not directly translate to the qualitative perspective. In qualitative research methods, such measures are replaced by the existence of trustworthiness, dependability, internal consistency, rigor and quality within the research data collection and analysis (Gasson, 2004). Triangulation is what helps the researcher to establish the existence of these qualities in her research.

3.1 Study Participants

A large urban public library system, located on the East Coast of the United States, allowed the solicitation of participants for this research during the various Friends of the Library monthly meetings. Participants were selected based on personal computer ownership, age, and an expressed interest in the collection and long term use of personal digital information. The study included participants ages 18 to 65. Purposive sampling was employed and a mix of gender and generations are represented in the study. The Drexel Institutional Review Board (IRB) gave approval for 35 participants; however, data saturation was reached after interviewing 26 participants. In a qualitative study data saturation is an indicator that data collection is complete, so the final participant group totaled 26 people.

As mentioned previously, public library users were specifically chosen as the focus of this study. Individuals who are primarily served by public libraries rather than academic or corporate libraries have access to similar information services and likely, but not certainly, exist in similar information environments. Further, in corporate and academic library environments institutional repositories are being developed to meet the long term digital preservation needs of scholars and corporate researchers (Lynch, 2003; Branin, 2005). No parallel services are being developed for public library users. While public libraries are involved in the development of digital collections, they are not involved with the implementation of institutional repositories as other types of libraries are. This research asserts that digital collections about local community events and history are the closest service to an institutional repository that public libraries offer. Given the difference in library services by type of library,

public library users are likely to have less knowledge of and access to resources related to digital preservation practices. As a result, public library users are underserved in terms of related services and research.

This library system and its patrons were specifically chosen for two reasons. First, it is a large urban public library system with tens of branches serving diverse communities throughout the city. The Friends groups selected for participant solicitation represent a range of social and economic communities.

Admittedly, individuals who are members of public library Friends groups do not necessarily represent the average public library users. These individuals are likely to be more civic-minded than the average public library user, and be more avid public library users than the average person. If this research were studying social phenomena such as civic engagement, social capital, neighborhood identity, then limiting study participants to only members of Friends groups would have the potential for creating bias in the data collected. However, the use and maintenance of personal information is more personal than social in nature and therefore should not be strongly influenced by social aptitudes.

Further, the additional characteristics required of the participants were of equal importance. Public library use is just one factor in four that made a participant suited to this study. The other three significant factors were: personal digital device ownership, an interest in collecting and preserving personal digital information, and age 18 to 65.

Recruiting participants for the study was accomplished at the Friend's meetings and with follow up emails and phone calls. People were very interested in

this research topic and were quick to talk about their own digital preservation disasters and their own lack of knowledge on the topic. Individuals asked to participate in the study because they were interested. Participant interviews lasted between 45 minutes and 2 hours, depending on how much the participant had to say. Many of the branch libraries had study rooms available for the interviews. Otherwise, local coffee shops were used to conduct the interviews.

While the Friends group members were diverse in terms of race, age, and gender, the participants were mainly women, white, and educated. Table 1 lists the participants in the order they were interviewed.[9] In sum, participants had the following characteristics: *gender*: 16 female, 10 male; *race*: 22 white, 2 black, 1 interracial Hispanic, 1 white Hispanic; *education*: 2 doctorates, 10 masters, 11 bachelors, 3 two years of college or less; *age*: 8 -20s, 4 -30s, 4 -40s, 6 -50s, and 4 -60s.

[9] Participant names are pseudonyms.

Table 1: Demographic Characteristics of Study Participants in order of Interview

Pseudonym	Gender	Age	Education Level	Hispanic	Race
Ann	F	61	Doctorate	No	White
Bill	M	25	Masters	Yes	White
Cathy	F	44	BS	No	White
David	M	35	BS	No	White
Eleanor	F	46	Masters	No	White
Frank	M	40	Masters	No	White
Greg	M	53	Some College	No	White
Holly	F	57	Masters	No	White
Irene	F	39	BA	No	White
Julie	F	65	BA	No	Black
Kelly	F	26	BA	Yes	Black, White, Native American
Lisa	F	50	BA	No	White
Mark	M	51	Masters	No	White
Nora	F	55	Masters	No	White
Oscar	M	40	Doctorate	No	White
Paula	F	20	In College	No	White
Quincy	M	28	Masters	No	Black
Robert	M	37	Masters	No	White
Sarah	F	62	BA	No	White
Tina	F	51	Masters	No	White
Uri	M	63	AA	No	White
Victoria	F	34	BA	No	White
Wendy	F	23	BA	No	White
Xavier	M	26	BA	No	White
Yvonne	F	23	BA	No	White
Zoey	F	24	Masters	No	White

3.2 Brief Overview of Methods

The data collection methods included the following: personal information matrices, semi-structured interviews, and a mapping technique. Data analysis includes the constant comparative method/data reduction, data displays, theory building and verification, and descriptive statistics of the transcribed interviews, information matrices, and information source horizon maps. Figure 1 provides an overview of the research methods within the context of the rest of the study design. Descriptions of data collection and analysis follow in detail.

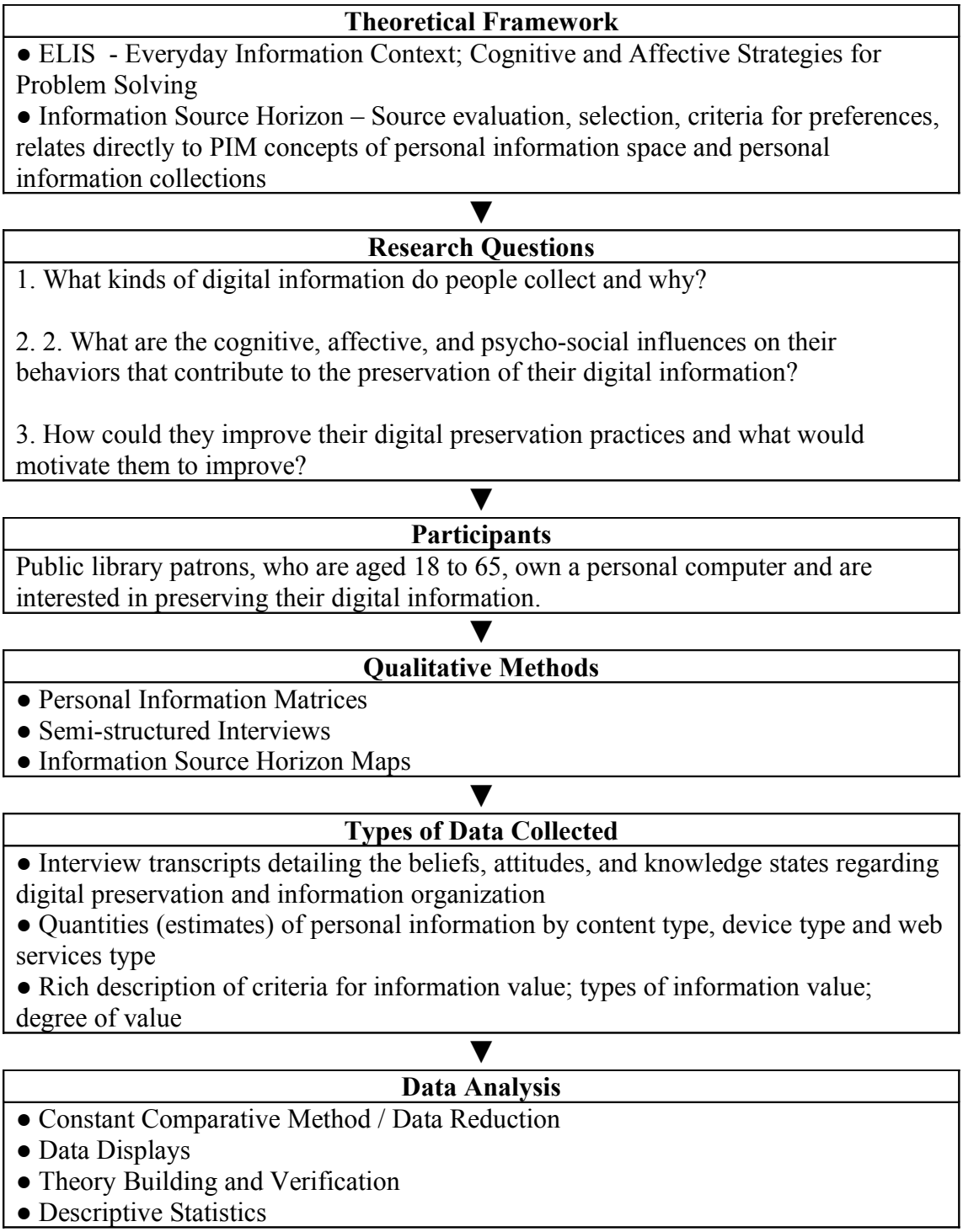


Figure 1: Overview of Methods

3.3 Evolution of the Research Design

This research took place in two phases. During the first phase the data collection instruments were adapted to better address the research questions. The first phase included four participants: Ann, Bill, Cathy and David. After analysis of the transcription of the first participant interview, it was observed that when the first participant talked about the value of her personal information, formats played a role in that discussion as this individual was transitioning from collecting physical or analog formats to predominately digital ones. Given this observation, Savolainen's and Kari's (2004) theory of information source horizons was adapted to my study to elicit criteria for how individuals valued their personal information. Information source horizon maps, described in sections 3.6 and 3.9 of this chapter, were used to collect data beginning with the second participant Bill. As the technique proved useful in addressing the research questions, the technique was used in all subsequent interviews.

During and after the first four interviews, a few questions were added to the interview guide. After the second interview, beginning with Cathy, all participants were asked if they treated content intended for long term use any differently than content intended for short term use. After the third interview, beginning with David, all participants were asked how they found their own content after saving it and asked to specify one piece of physical and one piece of digital content of great value. After the fourth interview, beginning with Eleanor, all participants completed the information matrices and were asked to describe any practices related to backing up and/or making redundant copies of personal digital files. Also, participants were asked what types of personal information they used most frequently. These changes

were made to get a more comprehensive understanding of the technical, cognitive and affective issues related to the preservation of personal information.

The matrices described in sections 3.4 and 3.7 of this chapter were added to the interview process to understand and to compare the technical information environments of the participants. The content types and storage locations used to create the two matrices were generated first through the identification of common content types, digital devices and web locations used by the average consumer. Then the matrices were sent to six different individuals to review and complete in view of their own personal computing/device usage. These six individuals varied in terms of occupational and personal use of computers and devices: one psychologist, one education researcher, two information studies PhD students, one home-maker and one computer programmer. These six individuals helped to identify the missing content types and storage locations and complete the fields of the matrices.

Ideally, counts of exactly how many each type of content stored by the participants would have been ascertained. However, the time involved to obtain specific counts would have been prohibitive. In a test of the matrices using actual counts of file types stored in each location, one individual spent three hours filling in the matrices. Also, the degree to which certainty of the counts could be achieved would be dependent on the individual participant's ability to make accurate assessments. For these reasons, it was decided not to compare participants using quantities of content.

The final 22 participants completed all three aspects of data collection. It should be noted that all interviews are used equally for theory building in the final analysis.

3.4 Personal Information Matrices – Data Collection

Two information representation techniques were employed in conjunction with semi-structured interviews. For the first of these two data gathering techniques, participants were asked to complete two personal information matrices. The matrices were designed to gather information on the types of digital content participants had stored in any digital devices and the various places of storage. The first matrix was designed to gather data on content types stored on various personal digital devices (laptop, cell phone, external hard drive). The second matrix is designed to gather data on content types stored on web services (YouTube, Yahoo, Snapfish). Please see Appendix A to see an example of the matrices.

The participants simply marked an X in each grid square that applied to the location of their content. By using the matrices, both the researcher and the participant were able to immediately analyze the distribution or concentration of the participant's entire body of stored personal information. While the participants were filling in the matrices they were asking questions and making comments. Both the representation of the participants' information environments and the conversation around the activity of filling in the matrices allowed the researcher to get a sense of how comfortable each participant was with different types of technology. In this way,

the matrices formed the basis for continued data collection though semi-structured interviews.

3.5 Interviews - Data Collection

As influenced by the multiple case study method proposed by Zach (2006), semi-structured in depth interviews were conducted with a select and small number of study participants taken from a larger population. See Appendix B to review the interview guide and Figure 2 below to examine the interview questions as they relate to the research questions. As is the practice within qualitative research perspective, the interviewing guide evolved and was influenced by the dialogue with the participants (Gorman & Clayton, 2005). Each interview varied slightly depending on the experiences of the participant but all 22 participants in the second phase were asked the questions as stated in the guide in Appendix B. Please see section 3.9 in this chapter for a discussion the two phases of the study.

The introductory questions were designed to elicit information regarding what kinds of physical and digital information the participants collected and any values associated with either format. The next questions were included to ascertain what steps were taken by the participants to preserve and organize digital information. Participants were asked specifically about their practices related to making redundant copies of their personal digital content. Also, included was a question regarding information use to explore the connection of frequency of use and value. The next set of questions focused on cognitive and affective issues. These questions asked participants to describe both a positive and a negative experience using digital

information. Also, participants were asked to discuss their preferences for information formats: physical or digital. With the ultimate set of questions, this research was seeking to understand what sources of information or human resources people turn to for technical information and how they prefer to learn technical information.

	Q1	Q2	Q3
I. Information General			
A. Complete Information Matrices	●		
B. Why do you save the information you do and how do you name your files?	●		
B. For information that you have in both physical and digital forms, do you value you one more than the other. Or just differently?		●	●
C. Describe one piece of personal information that is physical in form and one piece of information that is digital that you value greatly? Why is each valuable to you? How do you care for those possessions? How long do you plan to keep each one?	●	●	●
II. Personal information			
A. What types of information do you use most frequently? Why? How do you access items you use most frequently? Where are they stored?	●	●	
B. Do you treat personal digital information that you would like to preserve over many years differently than personal digital information you plan to only use for the short term? If so how so?	●		●
C. Do you back up all the content on your computer and/or digital devices?	●		●
D. What other steps do you take to preserve your digital information over the long term?	●		●
E. How do you find content that you have saved on your computer or other digital devices?			●
III. Positive Incident			
A. Please describe a positive experience with a digital information item, an experience that was joyful, creative or memorable?	●		
B. Start by discussing how you came to have the digital item in the first place? Why you decided to save or create it?	●	●	
C. What format did you save it in and why? Do you believe this format to be a good one for preserving this item over the long term?			●
D. Are you pleased with the format? Why or why not?			●
IV. Negative Incident			
A. Please describe a negative experience with digital information, perhaps you lost a digital photo or document?	●		●
B. How did this loss or experience make you feel?		●	
C. Since this happened have you changed the way you save these kinds of items? If so, how?	●	●	●
D. Thinking generally about using your computer and other digital devices to view or hear digital possessions, do you find that having to experience digital content via a computer screen and/or by some software application enhances or detracts from your enjoyment of digital content? Ask for experiences.	●	●	
V. Knowledge Acquisition			
A. Think of a time recently when you needed information regarding information storage, preservation, or information management. What type of information did you need and why did you need it? Who did you go to for help or what resources did you use? Were you satisfied with the information you received. Would you do the same the next time you needed this type of information or would you do something differently?			●
B. What is the best way for you to learn about computer or digital technology?		●	●
VI. Information Source Horizon Maps			
Two maps: information valued across all situations & information valued long term		●	

Figure 2: Interview Techniques and Questions as related to Research Questions

3.6 Information Source Horizon Maps – Data Collection

After the participants completed the interview, they were given two blank maps to fill in the three zones of their information source horizons. Directions are provided to participants in both written and verbal formats. See Appendix C for the written instructions and Appendix D to review examples of the maps. In the first map, participants were asked to consider personal physical and digital information that is of value to them across all situations in everyday life. In the center zone, participants were to indicate the information of most value to them in their daily life**.** In the middle zone, they indicated the next most valuable, and in the outermost zone the next most. In other words, the information should decrease in value as they move towards the outer zone.

In the second map, participants were asked to consider personal information they would like to preserve for at least five years, but as long as 25 years and beyond. With the most valued information items clearly identified, participants were then asked to detail the criteria they used for determining that value and how those were items stored and maintained. This representation of personal information helped participants to begin to make selections and distinctions within their own information collections based on a personal perception of value.

3.7 Personal Information Matrices – Data Analysis

The personal information matrices were analyzed in terms of the types and amounts of content the participants stored on digital devices and in web locations. This analysis was done in an effort to understand participants' personal information environments and to compare the information environments of all participants. The types and amounts of digital devices and web services that contain and provide access to personal information have implications for long term information organization, preservation and an individual's ability to curate her own personal information.

After the completion of the information source horizon maps, the participants were asked during the interviews to identify the location or locations of their most valued digital information, using their completed matrices to prod their thinking.

3.8 Interviews – Data Analysis

Analysis and theory building began with the first interview and evolved throughout data collection, which continued through July 2008. Data analysis entailed the use of the constant comparison method (Glaser & Strauss, 1967) to create an initial coding scheme. The constant comparative method is the most common method used for analyzing qualitative data (Agosto & Hughes-Hassell, 2005). Iterative comparison of emerging codes from the raw data eventually led to the derivation of emerging codes which were continually tested and refined. Concepts and theories developed from this research are emergent as yet no theory exists to explain the behaviors and affective responses that exist regarding the preservation of personal digital information.

During this process, interviews were transcribed and then read over several times all the while noting observations in the data in the form of codes and memos. Codes are labels used to describe or represent specific instances within the data. Coding analysis permits the reduction of vast amounts data in meaningful ways (Miles & Huberman, 1994). Codes organize the data and facilitate analysis through perceivable patterns and themes that result across coded data. Memos are conceptual notions about the data that tie elements of the data together and aid in the theory development (Miles & Huberman, 1994). Along with data reduction through coding and pattern identification, data displays are used to analyze and reduce the data to an accessible and meaningful order (Miles & Huberman, 1994). Data display is an analysis activity that facilitates understanding through the visual organization of data that allows for immediate access to conclusions made regarding the data.

To refine the initial coding scheme that emerged, the transcripts from the interviews were divided into 15 thematic sections. This allowed for easier comparison across participants as well as a more focused approach to analyzing each question. Codes were reviewed and revised for each question starting with the first response and then added to or modified in response to analyzing each subsequent participant response to that question. This process was repeated, modifying the coding scheme throughout the analysis of all interview questions. NVivo 7.0, a software program designed to aid the qualitative researcher, was used during the analysis to organize the coded text.

The coding scheme was evaluated using a measure of inter-coder reliability. The entire transcript corpus was coded first by the dissertation researcher and then a smaller subset of interview responses, approximately ten percent of the data, from five participants diverse in age and gender were selected and coded by two other coders for coding consistency checks. The two coders were PhD candidates in information studies trained in qualitative research methods and specifically in grounded theory. The inter-coder reliability measurement of agreement was 81% and 89% which are both greater than the established benchmark of 80% (Holsti, 1969).

3.9 Information Source Horizon Maps – Data Analysis

The information source horizon maps were analyzed in terms of what types of information items the participants valued and why. In terms of this research, value is defined as "what a person or group of people or group of people consider important in life" (Friedman, Kahn, & Borning, 2006, p. 349). The zones illustrate and contrast the types of items valued most for everyday use and long term preservation. Appendix D shows the maps created by two different participants.

The relationship that exists between physical and digital personal information items was measured using Savolainen & Kari's (2004) system of weighting each item in accordance with its location on the maps. Items in the center zone received 3 points, the next most outer zone 2 points, and then next most 1. The items closest to the participant on their horizon are given the greatest weight because the position indicates greater value. In addition to physical and digital formats, two other formats were indicated as valuable by participants: verbal and mental. Verbal refers to spoken

conservation with other human beings and mental refers to information which is stored in an individual's memory.

The descriptions provided by participants regarding why each item within each zone was valuable to them were analyzed using the same point system described above. Through the use of the information source horizon technique, the participants provided another perspective on digital items of value other than that provided during the interview. In combination, the two data gathering techniques provide a triangulated approach to understanding value as it relates to personal digital information.

3.10 Limitations

One limitation of this study is that findings are not fully generalizable to the larger population of U.S. public library users. The goal of qualitative research, however, is not generalizability, but transferability (Lincoln & Guba, 1985, p. 297). Qualitative research seeks to create a very detailed description of characteristics of a smaller pool of participants than quantitative research typically employs. Transferability indicates that the findings can do much to describe the behaviors of a similar pool of people, but it is understood that there are always variances in human behavior. Thus, the results of this study are intended to be largely transferable to the broader population of public library users, ages 18 to 65, but there will be some variance as contexts vary.

This research seeks to understand information behaviors of the individual in the context of everyday life and has chosen public library users as a primary function

of public libraries is to provide resources and services in support of everyday life and for other reasons previously mentioned. Given this, public library users are likely to be more aware of information resources and services in general than are non-public library users. This is a limitation of the study in that this awareness is likely but not certain to make public library users more cognizant of issues related to information organization and management as it relates to everyday information uses than are non-public library users.

Lastly, the extent to which an individual participant's personal information environment can be quantified and comprehensively described is limited by the inability to examine all their personal digital content stored across all personal devices and networked services. The estimates of amounts and descriptions of content and storage locations provided by participants are intended to serve as guide to understanding participants' environments rather than a definitive measure.

4. Findings and Discussion: Part 1: Technical Environment

The findings and discussion will be presented in together in the next three chapters with each focusing on one main aspect of personal digital preservation practice: technical environment, cognitive and affective factors, and social influences. As the aspects are integrated each chapter will touch on all three aspects to a certain degree. In addition to social influences, the third chapter will present an integrated framework for understanding the behaviors that contribute to digital preservation. To begin, this chapter presents the study findings and discussion on the influences of the technical environment that supports personal digital information, particularly examining storage devices used in general, for redundant storage, and to support digital information desired most for long term access.

4.1 Storage of Personal Digital Information (N=22)

All 22 participants completing the matrices stored more content types (range 6 to 18) on more types of personal digital devices (range 3 to 10) than they did on web locations (respectively: range 0 to 11; range 0 to 5). Digital devices were used to store content four to one when compared to web storage locations. (See Tables 2 - 3 and Figure 3. Appendices E and F provide data for each participant) Also, they stored more content on multipurpose devices, for example desk tops, lap tops, PDAs, than on devices designed only for storage. (See Figure 4.)

This is likely the case because participants need to save content to the device used to create it or lose the content. This is not always the case as one participant, Mark, indicated that he never saves content to anything other than an external hard drive. Mark has been using computers in the work place for 30 years and therefore he

has considerable experience using them. This creates more flexibility in terms file size and accessing his content from multiple devices. Also, this practice increases access as content saved to an external hard drive is more portable than content saved to a desk top.

(Mark) I probably have 12 external hard drives, you know, anywhere from 250 gig to like a tera.

(I) Now, why do you do that? Why do you not use the hard drive on your PC?

(Mark) Because I use, at any given time, I have five different computers that I work with, depending on the application and the programs that are on them. And so, it's a lot easier just to move…to plug in from one computer to the other than it is to try to carry the five computers around.

(I) So, you have a way to unify your collection? That's very smart.

(Mark) Thank you. But you know, you run programs and sometimes you run programs that, you know, potentially could crash a machine. So, I build the machine specifically with the applications in mind.

(I) That's fantastic.

(Mark) You have to have lots of accessible memory if you are working with video files and you want to do editing. You've got to be able to store a lot of chunks of that and so you can't really have too much else going on, on that machine.

His practice can be seen as a means of bringing the benefits of cloud computing to client side computing. Personal information storage is gradually moving away from a strictly client-side approach utilizing hard drives and local installations of software applications toward web-based storage services offered by private companies such as Google and Flickr, part of the cloud. Carr (2008, p. 113) provides this description of cloud computing: "….computing, as we experience it today, no

longer takes a fixed, concrete form. It occurs in the Internet's ever-shifting "cloud" of data, software, and devices. Our personal computer, not to mention our BlackBerry, our mobile phone, our gaming console, and any other networked gadget we use, is just another molecule of the cloud, another node in the vast computing network."

Along these lines, for the participants as a whole most of the content saved to the web were to subscription storage spaces, such as spaces offered by Mac.com or iDrive.[10] See Figure 4. By having subscription storage, the aspect of privacy associated with storing content on client side devices is maintained. Although, it is debatable whether privacy can ever be completely assured of information stored online.

4.2 Email and Social Networking Sites as Storage Spaces

Facebook and email are examples of storage locations influenced by social connectivity as well as technical connectivity and information storage. For example, Frank and Robert both lost all of their digital content due to a fatal hard drive crash and both were able to get some of their content back because they had shared their content using email. In the passage below, Robert recounts how he was able to retrieve some of his more important content:

(Robert) It was more work over the following, you know, over the following weeks, because when I would have to go and do another site visit, instead of like using a template I'd used before I had to start fresh. So, it was a little bit of extra work, but what I did was I went to my email and a lot of the documents that were the most important I had emailed to other people, so it was in my sent items folder. And then I would

[10] http://www.idrive.com/

> ask other people, say hey (people on my team that I work with) hey, email me, you know, the following document. So, I was able to reconstitute some of the more important stuff.

Content stored unintentionally in email highlights another benefit of cloud computing over client side computing: content is available regardless of the device being used. Facebook, like email, is not foremost a storage location, but four of the participants under the age of 30 (Kelly, Wendy, Yvonne, and Zoey), were unintentionally storing photographs as a result of sharing photographs through social networking.

Email is also being used intentionally to store content as indicated by these four participants below:

(Zoey) Just to make sure it's somewhere else. Like, if I can't back something up I'll email it to myself to have it somewhere and then probably back it up when I get home.

(Yvonne) I don't know I guess I feel that if I e-mail it to a few different accounts or something it is there and if I can't get into one e-mail account then I can get into the other one or something like that. It's nice knowing that the information is in a few different places.

(I) How do you care for your mystery novel in the making?

(Irene) Not particularly well, I'll tell you that much. When I got computer trouble, before taking it in for service, I emailed it to myself as an attachment. I have not printed it out. It is sitting on the computer.

(Wendy) probably because I can access it from anywhere, I feel like if it is not localized to one physical device it's less risky theoretically

4.3 Born Digital Content

Digital content stored by individuals is mostly content that is born digitally. Only a few participants had digitized content: scanned documents (N=4), scanned photographs (N=2) and scanned objects (N=1). See Figure 3. This is in contrast to libraries and other cultural institutions which are primarily dealing with digitized content (Dalbello, 2004). As mentioned previously in Chapter 2, academic and special libraries have begun to address the needs of their constituents related to the preservation of their born digital information through the development of institutional repositories.

Because of this difference, libraries have more control over their digital collections than do these participants who rely on a variety of digital devices and Microsoft products to create content. See Figure 4 and Table 42 in Chapter 6 for examples. As a result, personal devices and software programs are used to create, store, and maintain content. Because of this integrated process, individuals are not likely to give thought to collection scope or item appraisal or even acknowledge that they are building personal collections. Academic archives are just beginning to deal with unintentionally formed personal digital collections as donors begin to donate computers and other digital devices along with paper and physical artifacts. The donation of Salmon Rushdie's computers and other digital devices to Emory University is the first author who composed his works in a networked technical environment to make such a donation. Others will no doubt follow. Through the comparison of these different types of digital collections, greater insights regarding

collection development, maintenance and preservation of digital information can be gained than looking at these collections separately.

Table 2: Content Stored in Digital Devices for all Participants

	Devices	Content Types	Number of Devices Used for Most Distributed Content
Range	3 to 10	6 to 18	2 to 8
Mean	6.5	10.7	5.2
Medium	6.5	10	5

Table 3: Content Stored on the Web for all Participants

	Web Locations	Content Types
Range	0 to 5	0 to 11
Mean	2.9	4
Medium	2.5	2

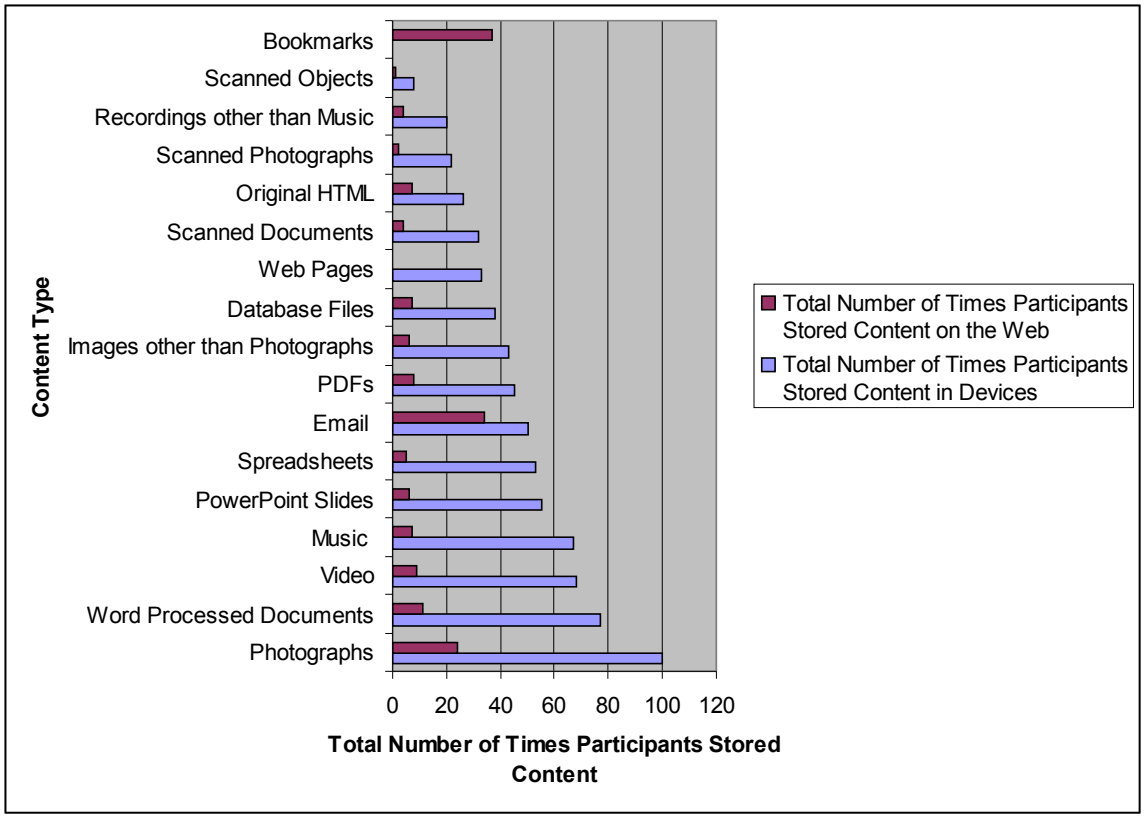


Figure 3: Comparison of Content Stored in Digital Devices and on the Web for all Participants

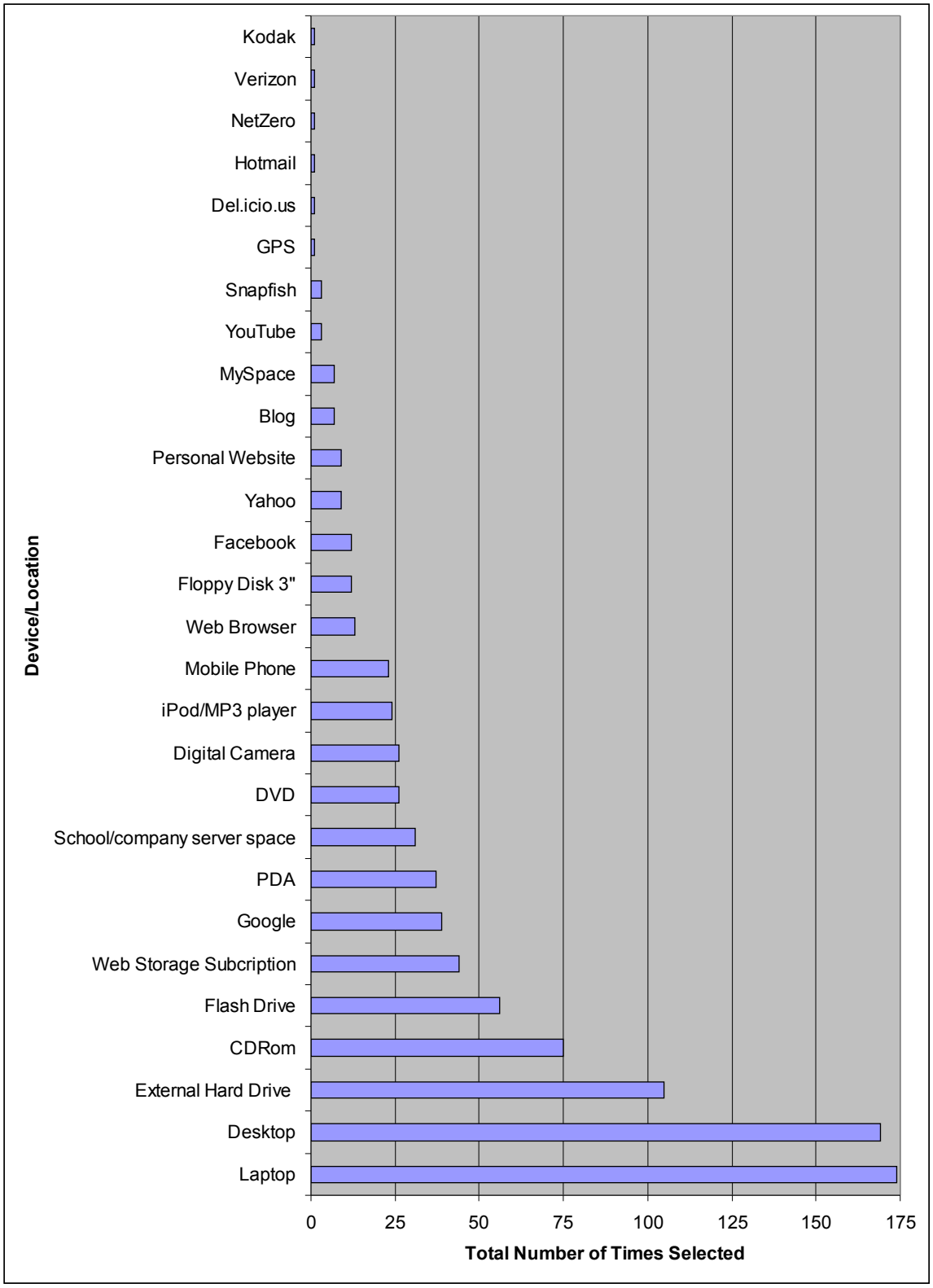


Figure 4: Storage Types for all Content in Order of Popularity for all Participants

4.4 Redundant Storage Decisions for Personal Digital Information (N=24)

When participants were asked if they treated digital content intended for long term use any differently than content needed only for short term use, initially 12 or 50% of participants indicated that they did. When pressed for how they did, 15 or 63% of the participants indicated that they made duplicate copies of selected files, files intended for long term use. Photographs were the most frequently duplicated item. Eight participants made copies of all their files and two of those eight routinely made images of their hard drives. (See Appendix G.)

Only seven of the participants routinely make duplicate copies of their digital content and one participant did not intentionally duplicate any of his content. For 17 or 71% of the participants, duplicating content is seen as a priority when prompted by anxiety rather than an everyday life routine. Lisa intends to transfer the burden of backing up her content to Verizon:

(I) So, do you just copied "My Documents" to the external hard drive?

(Lisa) Yes.

(I) And, how do you do that?

(Lisa) Um, when I get worried enough about it.

(I) Okay.

(Lisa) Not very often, you know, and that's what I need to do is get scheduled with the offsite back up. That's my ultimate goal, is to do an offsite back up with Verizon.

For Bill, the need to duplicate his content has become complicated by the use of multiple devices. He decided in favor of not dealing with his situation:

(Bill) Yeah, I have a desk top and a lap top , so part of the problem is that the files are mixed up, some of the files on the desk top are not on the lap top, I think there is software now that based on things that have been saved, to sync up the two. I really want to synch both of them up, and then take the conglomeration of the files and put them on the hard drive. I have only backed up a couple of times on the external hard drive,

(I) When was the last time you did so?

(Bill) Well my external hard drive kind of died, I have to get a new one, it has been my intention. I just haven't done it. It is my intention to do that, but when I did have one it was once every three months.

Automatic back up systems would be ideal if individuals did not use a multitude of devices and storage locations. No such technology exists to unify personal digital content stored on a diversity of devices and web locations that would facilitate comprehensive access to one's own content.

The occurrence of the four women who have their husbands duplicate their digital content for them indicates that maintaining personal digital content has the potential to become a household management issue. (See Appendix H.) However while convenient for these women, this practice potentially adds another layer between them and their content. Four of the seven participants who made redundant copies of their files were male and one of the females had her content duplicated by her husband. This could indicate that men are more aware of the need to systematically duplicate content. Further study would be needed.

Most participants, 10 or 42%, copied their content to at least two devices/locations other than the original location, six participants copied content to one other device/location, five participants to three other devices/locations, one participant to four other locations, and only one individual did not intentionally make redundant copies of any of his content. Overall most content was copied from a lap top or desk top to an external hard drive, seven and five participants respectively. (See Table 4 and Appendix H.)

Table 4: Content Copied from Initial Storage Source to Additional Sources for All Participants

Copied From ▫ Copied To ▫	Lap Top	Desk Top	External Hard Drive	PDA	Email	Totals
External Hard Drive	7	5	1	0	1	14
Flash Drive	5	4	1	0	0	10
CD-Rom	3	4	1	0	0	8
Web Storage	3	3	1	0	0	7
Email	4	1	0	0	0	5
DVD	1	2	1	0	0	4
Paper	2	1	0	0	0	3
Lap Top	0	0	0	2	1	3
Desk Top	1	1	0	0	1	3
Zip Disc	1	1	0	0	0	2
	27	22	5	2	3	

4.5 Storage Devices/Locations for Information Most Valued for the Long Term (N=22)

Participants most often stored digital information valued for the long term on digital devices that serve multiple purposes (N=17 or 71%), devices that can be used to create, share, view, and store content. Interestingly, the web also serves multiple purposes yet only six participants used them, and only seven content types were

stored there as compared to the 13 types stored on digital devices. As mentioned earlier, this contrast is likely due to the fact that content is most often stored where it is created. (See Table 5.)

Table 5: Most Valued Digital Information for all Devices and Locations for all Participants

Storage Medium □ Number of Items, Devices, and Participant for each Location □	Multi-purpose Devices	Storage Devices	Paper	Web
Number of Item Types Stored	13	12	8	7
Number of Participants That Use	17	13	5	6
Number of Types of Locations Used	5	5	1	4

Most of the participants' digital information desired for long term use is stored in at least one stable device; for example, desk top computers (N=9 or 38%), lap top computers (N=9 or 38%) and external hard drives (N=8 or 33%). Participants varied in terms of number of devices/locations used for their content. (See Appendix I.) For example, Frank, who experienced a total hard drive failure and lost much of his own creative writing, now stores his creative writing in three digital devices (lap top, external hard drive, and flash drive), a subscription web storage service, and prints out his writings to paper.

(Frank) All my digital files I back up weekly to an external hard drive. Also when I'm working on them I back them up to, well let's see I have a process for this, I back up to a flash drive regularly at the end of the day I back it up to an offsite tool, the server that hosts my web site and at the end of the week I back up everything to a hard drive. And that is a lesson hard learned because of a fatal hard drive crash on my lap top about 3 years that ate up some irreplaceable writing that I had down, stories, some fiction that I was working on that were getting ready for publication. One of those lessons you don't have to learn twice.

In contrast, Robert has some of his photographs stored on his desk top and some on the proprietary site owned by Kodak and he's not certain which ones are where or which photographs are in both places:

(I) Was there one particular piece of information? Is there one photograph or one email that stands out as one you would rather not lose?

(Robert) Mmm…yeah. There's one picture that I really, really like of our first child looking through a window and it was just an interesting picture I took. I had a digital camera at the time and the way that it…it just looked very nostalgic from the day that I took it. I think it's on O-photo now, and I keep saying I'm going to go back to my documents and…I'm sure it's on our computer at home too, I just have no idea where to look for it. But, I know it's saved in two places and for some reason that was a photograph that I have always really liked.

Ideally individuals would have access to information that details the benefits of one storage device over another. This would give them a sense of how reliable and appropriate a certain storage device/location is for different types of content. Digital video and music files in particular take up more space than most digital devices can afford. For example, Quincy who creates music using digital software has no back up for this content which is of great value to him. Because of space issues he only has this content stored to an external hard drive and no where else. If he had known to invest in web storage space in addition to the external hard drive, he would have copies of his musical creations and plenty of room for growth.

(I) Okay. Now, the music that's on the external hard drive but not on the hard drive of your desk top - Is that backed up?

(Quincy) No.

(I) No. So the unique content on the external hard drive is not backed up.

(Quincy) No.

(I) I suggest you find another hard drive.

(Quincy) Yes. You're right. You're right.

(I) Does your most current work, only exists on the external hard drive?

(Quincy) It does. And, see, all of this is so new to me that I don't get to complete the entire project as quickly as I would like to. Meaning, when I got the external hard drive, someone was able to walk me through the steps of actually putting it together and hooking it up, and I'm not that computer savvy.

I know music, but the technology piece of it just baffles me. I just don't really understand it. And so, I was able to clean up my hard drive, take off some of the things and save it to the external, but wasn't really sure how to just make it a seamless process, where I'm saving and stuff is saved onto the hard drive itself but being backed up here…

and I don't know if I had enough space to do that, so I just started saving here, on the external hard drive.

(I) Right.

(Quincy) But it makes sense, because, as you said, the unique content that's there isn't being backed up anywhere.

(I) Yeah.

(Quincy) So I need a back up for my back up.

(I) Exactly. So, this sounds like very valuable information and that it would be a big blow if something were to happen.

(Quincy) I agree with that.

(I) Because that's your own music you've created.

(Quincy) Yeah, it is.

(I) Okay. So, if you take away one thing today - go and back that up right away.

(Quincy) Dually noted.

4.6 Discussion

The participants' main digital preservation strategy is to make duplicate copies. While duplicating content ensures access in the event of a system failure, the resulting distribution of valuable content has the potential to decrease access as there is no unifying mechanism. Also, making copies of valuable content will not be enough to ensure long term access. The problems encountered with upgrades to operating system or software programs (See Chapter 5, Table 28) is a harbinger of other problems to come related to evolution of systems and file formats.

A study which explores the longevity of personal information collections is needed to identify the problems likely to be encountered. Existing research indicates that storage media is likely to present the smallest obstacle to access over the long term and the issues related to technology obsolescence will present the greatest hindrance to access (National Diet Library, 2008).

5. Findings/Discussion: Part 2: Cognitive and Affective Issues

In this chapter, cognitive and affective issues related to the preservation of personal digital information are explored. The ways in which individuals think and feel about their personal digital information has the potential to influence their long term access to that information. Understanding why and how individuals save and use information can help researchers design better systems and methods of teaching to support and improve personal digital information management. This chapter explores participant behaviors related to saving, organizing, and re-finding content; format (physical versus digital) preferences; value and formats; the ability to specify personal information of great value; information use as a predictor of value, and negative affect associated with information management.

5.1 Saving and Naming Files (N=26)

Most of the participants saved digital files to access them later if needed (N=15 or 58 %). (See Table 6.) However, as all participants (N=26) created file names inspired by the current context rather than long term access there appears a mismatch between the reason for saving the content and file naming practices. (See Table 7.) The following comments from Victoria, Paula, and Lisa illustrate this point:

(Victoria) If it's a file that I've been working on, on my desktop, which is where I do most of my word processing and other things, I first save it to my hard drive.

(I) When you name the file – how do you name it?

(Victoria) If I type up a letter for my father it would be "letter to" whoever it is, possibly a date and then it's saved in the folder under "Dad", for instance.

(Paula) It depends, if it's a really important like a school document, I save it with something like a good name but if it's something stupid, I just bang the computer.

(I) Give me an example of a good name.

(Paula) A good name would be like "Costco Final"

(Lisa) I name it according to its content.

(I) Can you give me an example?

(Lisa) Like, "kids at the beach".

(I) Okay.

(Lisa) And then I will put the year.

(I) And then that goes into a folder kind of thing?

(Lisa) Yeah. Well, I don't know if it's in a folder right now. I think it's just in My Pictures. Like, within My Documents there's My Pictures.

This finding supports a similar finding by Barreau (1995), The finding that participants created file names based on the context of the current information use is supported by the same finding in Barreau's (1995) research of seven office managers and their personal information practices. In addition to current context naming protocols, no technical system or software application exists to unify all files saved across multiple applications and devices. A program that could do this would have the potential to provide some degree of continuity over time. None of the participants reported creating their own system to unify their content. Jones (2004) has suggested that individuals create a personal unifying taxonomy to provide this kind of organization. This labor-intensive task seems to go against the convenience deemed

by the participants as inherent to digital information, which is discuss in the next section of this chapter.

Also, most consumer software does not allow for metadata other than file names. Digital cameras assign a number to each file they create thereby reducing the need for or the chance that metadata will be applied at the item level for photographs. Folders are relied upon to apply an overarching description to a group of digital photographs which sometimes number in the hundreds. Study participants applied files names to only a limited number of individual photographs but assigned a general description to a folder containing photographs from one event. Many participants described their process much in way Sarah has described hers:

(Sarah) Well, I have a structure of folders that works for me. I keep them under My Documents for the most part, and then the photos go under My Pictures with subfolders for…for instance under My Pictures I have "Family and Friends" and then under that I have photos…usually the subfolders are named by the date. Somewhere in the name of the folder I have the date that they were taken. If it was a trip I'll say "Sarah's Trip 2006". And then I have other categories, like my house here or if I take a trip I'll just have a "Trip 2005" let's say, and then underneath that I'll even have subfolders for if I need to, my different locations on my trip.

Below, Irene indicates that she sometimes uses the file name supplied by the camera and sometimes she will apply file names to her pictures in an effort to remember what the photograph represents. Overtime specific memories are incorporate into existing schema which provide mental framework for understanding new information and for remembering memories. The details of specific details of memory may be altered or mistaken based on one's own schema (Bartlett, 1932 &1958; Rumelhart, 1980). This understanding would support the need for individuals

to use more detailed metadata schemes when classifying their own information. That the current context used to names files will not be accessible by an individual's long term memory with any specificity but rather the memories recalled will be generalized in accordance with the individual's mental framework.

(I) What about a picture of one of your children?

(Irene) I have a lot so sometimes I leave them as whatever the default is. But usually it will be a name, if I do name it, something that, the name and activity, something to jog my memory so I can remember.

5.2 Organizing Files

Photos are the most collected and most distributed content. (See Figure 3 and Appendix E.) It is likely that photographs maybe the most difficult to re-find because the reliance of folders for describing large amounts of content and the absence of searchable text. Ann believes the fate of most digital photographs will land in the "digital ditch."

(Ann) The digital form is more immediate but because I haven't catalogued pictures that I took 5 years previous. Shoot something on the weekend, I will send out the pictures that I like, especially if I'm in it and I look good. I will send those out right away and then they go into the digital ditch where as the ones that I have picked up from the drug store or if I really care about the film from a serious processor. I have picked up taken home gone through, edited, glued them or put little black corners on them put them in an album with a series of photographs, so I've done things with them. The immediacy of the digital diminishes the importance of it in the long term.

Photographs that are shared are likely to have greater longevity than photographs not shared as they have been selected for their worth and they are now likely to be maintained by two or more individuals. Photographs sent to others using

email might be easier to re-find using the search mechanism of email programs as the email message delivering the photograph/s is likely to acknowledge the reason for sending the photograph.

In general, folders are relied on heavily by the participants to re-find content. Likely this is due to the limited opportunity provided by personal computing systems to create metadata for individual files, which in turn limits the robustness of search functionality. This offers a sharp contrast to how information is found on the web which is primarily through search. The ability to search metadata has the potential to enhance one's ability to access personal content over the long term. This presents an advantage of web based storage locations over digital devices. Re-finding specific photographs will likely present problems in either environment, at least until image-based content searching is perfected (Gonzales & Woods, 2007).

Table 6: Reasons for Saving Digital Information for all Participants

Reasons for Saving Digital Information	Number of Participants
Future Reference	15
Connection to Self	8
Increased Access	5
Emotional	5
Born digital	4
Sharing	4
Evidence	4
Represent Memories	3
Investment of Time	3
Save Everything	3
Family History	2
Facilitates Editing	2
Work In Progress	2
Use Everyday	2
Easy to Organize	2
Paid for It	1
Private	1

More than half of the participants (N=14 or 54%) developed a personal but informal classification system for their files systematically using dates, special naming practices, and complex folder labeling. However, this process too appeared to be largely motivated by the current context with no documentation or system unifying the description and location of their personal digital content to support long term access.

Frank has three facets to his personal document storage system: work, personal, and web related. While this helps to keep broad aspects of his life organized, as he admits, there is some cross over between work and personal. Other comments he made during the interview indicated that the two websites he maintains relate to both his work and personal interests. Because of the broad nature these categories, overlap is guaranteed. It is likely knowledge of his current information needs and uses that provides the context necessary to access specific information.

(Frank) I don't use the existing the default file structure that computers default to, I have a big folder called Frank's stuff, within that I have the work, personal, and web related. Three major sub directories, all the personal stuff goes in one, all the work related files go in another, that includes my own writing my business the projects that I get paid for and the projects that I right for myself that I sent out hopefully get published or paid that I'm not doing for anyone else, that's where the cross over happens. And the third one is web related. All my HTML files, photos, and the two web sites that I maintain for myself.

Nora's practice of including her initials in the file names of content she's created to differentiate it from content created by others. This is a sound practice for anyone who frequently exchanges digital files with others. A simple search on her initials can produce all the files she's created. Further study could investigate the

extent to which personal information has been personally created and to what extend it was been collected from other sources.

(Nora) I usually name it by something that if I look at it I can reference it. If I'm doing a flyer or something I'll put "flyer Springfield library". You know, abbreviated. Or, I'll put "letter…" whatever. And, all of my files, because it's a Mac, I put it as NWK.

(I) Which is?

(Nora) It's just a little acronym that shows that it's mine and not Mac. So, I can differentiate them from other things people have sent me and things that have been converted from PCs. You know, so I know which are my original documents and which ones I looked over and sent someplace or whatever.

Tina relies heavily on dates and folders to organize her content. She only describes her content at the folder level and not at the level of individual items. This is no doubt an efficient and meaningful system to her; however, if for some reason the content is separated from its folder, then future access to the content could become a challenge. In time, individuals may find that they need to migrate their content from one operating system to another, or one program to another, and the folders they rely on to categorize and re-find personal information could be lost in the process. It is for this reason that it is best practice to store metadata and digital object together in one file. See the discussion of the OAIS Reference Model in Chapter 3.

(Tina) Let's see…I want to make sure the file and the folder show up chronologically correct. So, rather than just going like "January, February, March", okay, I don't do that. Because February comes before January alphabetically. So, I would do…and I don't do "1, 2, 3, 4" because 10, 11, and 12 are two digits. So I do 01 for 01 to really show up as 1, the first one.

(I) Okay.

(Tina) So, I do 01 and then I may put JAN, which is still going to show up as number 1 anyway. So, probably the…and I just follow that logic through whether it's a file or folder.

(I) And then do you ever include something that would reflect the content of the item you're saving?

(Tina) Probably at a folder level. So, for example, I would go "2008" and then I would go 01 JAN and then maybe "spring wear fall" or maybe "garden project" or you know "front porch" and sort of categorize it that way.

(I) And that level of categorization is to the folder, not to the file?

(Tina) Correct.

5.3 Re-finding Files (N=23)

Half the participants (N=13 or 57%) seemed to understand that how they organized their information was important to re-finding their own information again. (See Table 7.) Holly and Xavier provide examples of this below. Holly has designed her organization with her family in mind so that not only she can re-find information but hopefully so her family can as well.

(Holly) Since I've watched my mother age and see what happens I try to stick to as much reality as possible. My daughter, her nickname is from the time she didn't move at all as a child, she's Slug and she's always been Slug and that is her. And yet everything is fairly obvious, recipes is under recipes, my French lessons are under French lessons. My son is under his first name. It's very easy, so if anyone wants to break in they can break in.

Xavier speaks to the need to create his own mechanism for version control in order to create efficiency in re-finding and re-use of personal information.

(Xavier) Okay. What I do to name a file is I generally choose something about the keywords that have to do with it and also maybe choose a date or something else that might be indicative not only of what it is, but when it was produced, because I may have several documents or files that have roughly the same content or the same content, but need to be differentiated from each other either based on the drafts or the version of it or when it was produced.

Most participants (N=21 or 91%) browsed existing folders to re-find information they saved rather than using a search tool. (See Table 8.) As mentioned earlier, the design of personal computing systems encourages browsing over search. Discussion from Eleanor and Wendy illustrate this point.

(Eleanor) My file structure is pretty good, I browse, the Apple has the little search thing called Spotlight which isn't bad, and if I can not remember where it might have gone, I will do a full text search for it, but I browse first.

(Wendy) I usually search the documents; I try to be careful about labeling folders.

(I) Do you search first?

(Wendy) When it's not obvious where it's located, or if it's old, or I can't remember which folder I put it in.

(I) Would you say you browse your folders more or you search more?

(Wendy) I probably browse more first and search as a last result.

Both Eleanor and Wendy indicated that they use the search when their memory has failed them or when their long term memory has not cued them to the right association. Again, this supports the need for more detailed metadata to support long term access to personal information.

Table 7: Organization Method for All Participants for all Participants

Organization Practice	Number of Participants
Context Motivated File Naming	26
Informal Classification System	14
Awareness of the Need to Organize for Long Term Access	13

Table 8: Method for Re-finding Content by Individual Participant

	Browse	Browse, Then Search	Search, Then Browse
David	X		
Eleanor		X	
Frank	X		
Greg	X		
Holly	X		
Irene		X	
Julie	X		
Kelly	X		
Lisa		X	
Mark	X		
Nora	X		
Oscar		X	
Paula	X		
Quincy	X		
Robert	X		
Sarah			X
Tina	X		
Uri	X		
Victoria			X
Wendy		X	
Xavier	X		
Yvonne	X		
Zoey	X		
Totals	16	5	2

5.4 Preference for Information Formats (N=26)

Participants were asked two different questions to determine the extent to which a preference for digital formats is evolving over physical ones. (See Tables 17 and 18.) Both questions yielded similar results but not always with the same response from each participant. Only five participants answered the two questions in similar ways. Two male and one female participant indicated a general preference for digital information regardless of information need or context and two female participants indicated a preference for physical formats. However, most participants (N=21or 81%) indicated that the format preference depended on the information need or use either by directly saying so or by indicating a preference for digital information in one question and digital in the other.

In the first question, participants were asked to compare content they had in both a physical and a digital format (e.g. photographs, address book etc) and to indicate if which format they preferred or if the formats were valued equally but differently. The participants' preference was expressed as a matter of personal taste and also a matter of information need. Photographs drew the sharpest contrast with regards to personal taste. In five of the seven cases, a preference for physical information, the item being compared was photographs. Kelly and Eleanor illustrate this preference below:

(Kelly) I think pictures I value more in the physical format.

(I) Why is that?

(Kelly) I don't know, I guess because I can hold it, but then again if it is digital, you can pick it up and take it as well, I guess I'm just an original old school type of girl in some ways

(Eleanor) with the photos, generally you take them, the whole kit and caboodle and pour them into your computer, the ones I value, I sent out to make a print of, so that I can look at it all the time, it is placed somewhere I can see it, I think books are the same way, its nice to look at yourself and see your favorite titles, you can't open up your lap top and see your favorite books, its partly tactile, lot of it has to do with ownership, and usually with the photos, better quality.

(I) What criteria would you attribute to the value of the physical items?

(Eleanor) It has to do with that they are immediately accessible as opposed to opening a computer, finding the directory, finding the particular file, clicking on it, opening up the particular file you are supposed to view it, its just faster.

Six of the seven instances a preference for digital photos expressed for reasons given included saves space and easier to find again when compared to printed photos. Julie expresses a preference for digital photographs for exactly the opposite reason that Eleanor favors them:

(Julie) Well, if it is in a physical format I'm always trying to figure out how to get it in a digital one. And back it up. Because even pictures now, they take up too much room, if I can store them digitally, and now with digital photos, and when company come over, just load them on a flash drive and show them the digital whatever, why do you need all these pictures and a bunch of photograph albums collecting dust? I just tend to always be looking at how to get rid of paper. I have a little library in my room, even books are getting so that they take up more and more room, so I am always looking for ways to part company with stuff that I've collected.

Because this question drew the participants to contrast digital and physical formats, they expressed values related to the characteristics they believed to be inherent to the medium.

(Irene) I think you have more respect for something that is actually made. They fulfill different uses. If I'm redoing my resume, I'm going to value the one on the computer more because that's one that is a lot easier for me to use. I have an easier time reading something that is printed out a manuscript or something like that. I can write an entire book on a computer. I would entirely value it more if it were printed out. You know what I mean. 50,000 words on a computer is perfectly fine. But if I printed it out, it looks like a book. It's more valuable to me because it looks like a book. I've had books my whole life, a book is a thing, books are in the library.

In expressing a preference for digital information, participants described the value of digital information in terms of the technology that supports that information rather than information itself. They spoke to technology saving them time, space and energy. Digital information is cheaper, easier to access, universally available, portable and immediately available. Whereas, with physical information, the values expressed described physical attributes, you can hold it, touch it, and write on it. Also they spoke to qualities believed to embody physical information: real, original, old fashioned, irreplaceable, and truthfulness.

Table 9: Indication of Preference for the Same Content that is Both Digital and Physical by Individual Participant and for all Participants

	Item/s Compared	Depends on Context of Information Need or Use	Preference for Digital	Preference for Physical
Ann	Photos			X
Bill	Letters/Email	X		
Cathy	Email		X	
David	Photos		X	
Eleanor	Photos/Books			X
Frank	Contact Information	X		
Greg	Financial Information/ Letters	X		
Holly	Recipes	X		
Irene	Resume/Books	X		
Julie	Photos		X	
Kelly	Photos			X
Lisa	Photos		X	
Mark	Calendar	X		
Nora	Personal Journal			X
Oscar	Legal/Music	X		
Paula	Photos			X
Quincy	Music		X	
Robert	Photos	X		
Sarah	Photos/Music	X		
Tina	Photos		X	
Uri	Photos		X	
Victoria	Calendar/Contact Information		X	
Wendy	Journal Articles			X
Xavier	Photos	X		
Yvonne	Photos		X	
Zoey	Books	X		
Totals		11	8	7

Most participants found (N=21 or 81%) that generally technology enhanced their experience of information but not always. Greg and Xavier provide examples of some why this is the case. They found experiencing information digitally enhanced their access to information and to other people. However, Greg indicates that digital information lacks the warmth of physical information received from other people. Xavier indicates that while digital information enhances many aspects of his everyday life he finds he needs a break from his digital life from time to time.

(Greg) Depends on the circumstances if I want to look at something and I don't have a computer then what good is all the technology in the world if you can't access it?

(I) Sure.

(Greg) Double edge sword really, in many ways it's a positive thing, instant access to so many different things, on the other hand there is the essence and the feeling of a handwritten letter that you are reading that isn't a cold printed out email, there's give and take on both.

(Xavier) It can be an enhancement in the fact that it makes information more readily available and makes people more accessible. I have family...none of my family lives around this area. They live all over parts of the country and, you know, while it's easy to use a cell phone or a land line to talk to someone on the phone, it's more difficult, like I said, to mail prints of pictures or CD's than it is to send them in email. And, also, reading news, it's possible to follow up on an article, and online learn more about that, or get a different perspective on what this person thought happened. So, I generally find that an enhancement. Other times, I find it kind of annoying how dependent I've become on digital information. There are days where I'll say to myself, you know, I'm just going to boycott all of this stuff, and I won't…I won't look at the computer. I won't use my cell phone. I won't turn on – we don't have cable or any sort of television at home – but just stay away all of that and I usually find that to be, you know, I don't really feel any negative consequences from that. By and large I think it is the positive enhancing experience.

Table 10: Indication of whether or not the Experience Information in a Digital Format is an Enhancement or a Distraction by Individual Participant and for all Participants

	Depends on Context of Information Need or Use	Enhances	Detracts
Ann	X		
Bill			X
Cathy	X		
David		X	
Eleanor		X	
Frank	X		
Greg	X		
Holly	X		
Irene			X
Julie		X	
Kelly	X		
Lisa			X
Mark		X	
Nora			X
Oscar	X		
Paula			X
Quincy	X		
Robert	X		
Sarah		X	
Tina	X		
Uri		X	
Victoria			X
Wendy		X	
Xavier	X		
Yvonne		X	
Zoey			X
Totals	11	8	7

5.5 Value and Personal Information Formats (N=25)

Savolainen's and Kari's (2004) weighted point system was used to calculate the value placed on format of personal information valued for everyday and over the long term. The point system is described in Chapter 3. Digital formats were valued more frequently by participants (N=21 or 84%) for everyday information needs and uses in comparison to physical formats (N=4 or 16%). Physical formats were valued more frequently by participants (N=16 or 64%) for long term information needs and uses in comparison to digital formats (N=8 or 32%). One individual valued physical and digital formats equally for the long term. (See Table 11.) When comparing the overall values expressed for both digital and physical formats, digital formats were valued 65% of the participants valued digital formats for everyday information and 49% of the participants for long term information. For both everyday and long term information needs digital formats were valued by 57% of the participants. (See Table 12.) While physical formats were more prominent on the long term map – they were just barely so. This indicates a move towards collecting more digital information of value and the need to improve the digital preservation practices of public library users.

Table 11: Format of Valuable Personal Content for Everyday and Long Information Source Horizon Maps by Individual Participant

	Everyday map totals					Long term map totals			
	Digital	Physical	Verbal	Mental		Digital	Physical	Verbal	Mental
Bill	17	14				11.5	6		
Cathy	8	47				0	18		
David	11	7				3	13		
Eleanor	5	1				1	5		
Frank	10	9				14	10		
Greg	4	2	1			3	11		
Holly	14	4				10	12		
Irene	6	4		6		7	12		
Julie	18	0				6	6		
Kelly	19	8				7	11		
Lisa	12	0				3	4		
Mark	7	10		8		4	1		2
Nora	12	18				13	20		
Oscar	29	0				30	0		
Paula	9	6				5	8		
Quincy	9	0				3	6	1	
Robert	17	0				5	15		
Sarah	14	13				18	20		
Tina	10	0				16	3		
Uri	17	2	2			3	0		
Victoria	9	13	5			0	14		
Wendy	14	1				12	17		
Xavier	14	0				22	0		
Yvonne	15	1				10	6		
Zoey	15	8				11	13		
Totals	315	168	8	14		217.5	231	1	2

Table 12: Format of Valuable Personal Content for Everyday and Long Information Source Horizon Maps for All Participants

	Digital	Physical	Mental	Verbal
Everyday	**315**	168	14	8
Long Term	217.5	**231**	2	1

5.6 Specificity in Identifying Items of Great Value (N=23)

During the interview, participants indicated one specific digital and one specific physical piece of information that was of great value to them. These are listed in detail in Appendix J. Then their responses were checked against the information source horizon maps to see if the participants included these items on their maps. (See Table 13.) Of the 46 specifically mentioned items only six were included on the maps: three digital and three physical and two for everyday context and five for long term (one item was mentioned on both maps). Participants were more inclined to include general content types rather than specific items as 12 or 48% of the participants included the general content type of the specific item rather than the specific item itself. Five or 20% of the participants did not include the specific items or the general content type that represents them, three provided both general and specific and 3 only specific. (See Tables 14 and 15.) This finding is supported by research that indicates long term memory privileges organization and categories to create efficiencies for individuals so that most of life becomes routine rather than a continual learning experience that would tax and overwhelm their short term or working memories (Rasmussen, 1990). Again, this provides support for the need for metadata especially when dealing with photographs which are inherently without searchable text.

Table 13: Information indicated as having Great Value that also appears on the Information Source Horizon Maps for all Participants

	Everyday				Long Term		
	Zone 1	Zone 2	Zone 3		Zone 1	Zone 2	Zone 3
Physical	—	—	1 General		8 General 3 Specific	—	—
Digital	3 General 2 Specific	3 General	1 General		8 General 2 Specific	2 General	—

Table 14: Specific Items Indicated as having Great Value that also appear on the Information Source Horizon Maps for all Participants

Specific Items of Great Value	
Digital	Physical
QuickBooks	International Adoption Documents
PDA	Wedding Rings
"Photo Library"	Wedding Photo Album

Table 15: General Content Types Indicated as having Great Value that also appears on the Information Source Horizon Maps for all Participants

General Content Types of Great Value	
Digital	Physical
Photos (9)	Photos (5)
Papers for School (4)	Books (2)
Voice Mail Messages (1)	Art (1)
Documents for Volunteer Work (1)	Furniture (1)
Financial Information (2)	

Participants readily identified one piece of physical information of great value to them. In contrast, every participant had difficulty identifying one particular digital item of great value. Either the participants thought of their digital information as collections of content rather than as unique items or they had trouble associating value with digital information.

(Zoey) Yeah. It's harder I guess to value digital things. Maybe I'll say like some online photo albums that I always think I'm going to order the prints, but I never do, so they are just online.

(I) Are you thinking of collections of pictures rather than one in particular?

(Zoey) Yes.

(I) Why is that?

(Zoey) Well, I guess I don't really have one in particular that I value above all others, but if I were to lose the whole collection that would be pretty traumatic.

(I) And one piece of digital information of great value to you?

(Holly) Thinking…. actually nothing. I value some of the pictures of the some of the things I've done.

(I) There's not one that you'd be saddened by the loss of?

(Holly) No, you can have the recipes, I'd be devastated by the loss of my PDA, because that would be the loss of its content and not any particular content but its general content.

(Bill) That's like my life, honestly, at the same time it's like I was saying before, most of that stuff I'm never going to look at it again unless it's sentimental, there is something very comforting to having all your files.

Like a lot of people freak out when they lose their stuff, when their computer goes bad and they lose all their files, the reality is most of those files didn't mean anything, they will never look at them again, its just difficult, it's a sentimental attachment.

What Bill has expressed regarding not using but valuing his digital content, he also expressed regarding physical content he keeps in something he calls his sentimental folder:

(I) Do you value your physical information more than your digital?

(Bill) I don't know I wouldn't differentiate it. I don't know that there is, I mean it is hard to tell, depending on the information, of course, I have emails that are very meaningful to me, just as meaning as some of the stuff that would be in the file, it is nice to be able to touch it, especially if someone has written it by hand. Honestly, that sentimental folder, I don't really every ever go in there.

In this way, personal information that is valued for sentimental reason is likely to be used in or thought of the way that Bill has expressed his use of this type of information above: valuable but not useful.

All participants (N=25) but one (N=1) could think of a digital information item or a collection of items that was valuable to them when pressed to do so. Physical items of great value typically pointed to others or were product or object oriented (N=6 or 24%). These items connected the participants to family (N=13 or 52%) or society (N=2 or 8%) and to specific life events (N=2 or 8%). For example, a coffee table passed down through generations, international adoption documents, a photograph of grandparents' wedding, a rare book, a painting inherited. When pressed to name a specific digital item of great value, participants typically identified items that represented their skills and abilities with technology and/or hours of hard work (N=12 or 48 %). Some of the processes they spoke of included creating a website, setting a photo story to music, and writing art history final. Also with regards to valuable digital information, the connection to self (N=11) was greater than to

connection to family (N=5 or 20%) which is in contrast to the valuable physical information discussed: connection to self (8 or 32%) and connection to family (13 or 52%). (See Table 16 and Appendix J.)

Table 16: Values Associated with Specific Physical and Digital Items for all Participants

Values	Physical	Digital
Connection to Family	13	5
Connection to Self	8	11
Documents Life Events	2	4
Historical	2	0
Monetary	4	0
Financial	0	2
Aesthetic/Creative	1	12
Tells a Story	2	1
Preference for Tangible	1	0
Irreplaceable	2	0
Saves Time	1	1
Sharing	1	0
Supports Work	0	1
Practicality of Digital Information	0	5
Future Reference	0	3
Connection to Friends	0	1

5.7 Use as Predictor of Information Value (N=22)

During the interview, participants indicated the types of information they used most frequently. Then their responses were checked against the information source horizon maps to explore use as a predictor of value. The data provide some support for use as a predictor of value, particularly for digital information valued in the everyday information context. (See Table 17.) Of the 31 information items mentioned by 22 participants, 27 or 87% of the items were digital. Of the 27 digital items indicated as most used, 18 of these items were indicated as valuable on the everyday

maps and 11 on the long term maps. Only two physical information items were indicated as most used and both appeared on the everyday map and one on the long term map. Verbal information was indicated as most used twice and appeared once on the everyday map. (See Table 18.)

Digital information was the most frequently used type of information. Use has the potential for predicting value and could prove helpful as people acquire increasingly large amounts of digital content. But many of the software programs and operating systems offered by Microsoft and Apple do not offer ways to monitor the usage frequency of personal files.

Table 17: Information indicated as Most Frequently Used and whether or not that Information is also indicated as Valuable on the Information Source Horizon Maps by Individual Participant

	Most Frequently Used Information	Everyday	Long Term
Eleanor	Web Pages		
	Email	X	
	Periodic Databases		
Frank	Web Pages		
	Discussion Boards		
Greg	Human Conversation	X	
Holly	Cell Phone	X	X
Irene	Web Pages		X
	Internet		
	Human Conversation		
Julie	Online News	X	
Kelly	Email	X	X
Lisa	Email	X	X
Mark	Email		
	Music		
Nora	Print Personal Journal	X	X
Oscar	Web Pages		
	Financial Information	X	X
	Email	X	X
Paula	Email	X	
Quincy	Cell Phone	X	
Robert	Email	X	
	Digital Calendar	X	
	Online News		
Sarah	Email	X	X
Tina	Email	X	X
Uri	Blogs	X	
Victoria	Print Newspapers	X	
	Digital Photographs	X	X
Xavier	Music	X	X
Yvonne	Digital Photographs	X	X
Zoey	Email	X	X
Totals		21	13

Table 18: Information Indicated as Most Frequently Used and whether or not that Information is also indicated as Valuable on the Information Source Horizon Maps for all Participants

Format	Number of Items	Valued for Everyday life	Valued for the Long Term
Digital	27	18	11
Physical	2	2	0
Verbal	2	1	0

5.8 Negative Experience associated with personal digital information (N=26)

Most of the participants' negative experiences involved problems with digital devices and loss of content, either a technology malfunction (N=12 or 46%) or a system incompatibility that resulted from upgrading operating systems (N=6 or 24%). User error accounts for the second most frequent negative experience (N=7 or 28%). Frustration (N=9 or 36%) was the emotion most associated with these experiences, followed by anger (N=7 or 28%). Given that most negative experiences involved loss of content, behavioral changes related to saving content in more places (N=6), more often (N=4) and more carefully (N=4 or 15%). Also, two participants decided to store content on the web because of losses due to devices malfunctioning. (See Tables 19-21.)

While loss of content inspired better participant practice with regards to duplicating files, Frank was the only participant of the 12 who lost content due to a hard drive failure that developed a many tiered system for duplicating important files. Part of his system involved storing content on a subscription web storage account, in case all other devices fail. Oscar is the other participant who moved his content to the web in order to have yet another location for his content independent of devices.

The fact that system failures and incompatibility issues are commonplace among personal digital device users likely contributes to the ambivalence about digital information. Values may not easily be attached to that which is considered ephemeral. Bill and Robert expressed this notion as the cost of doing business, as in a parking ticket or a flat tire:

(Bill) I have lost some papers, nothing sentimental, I'll be working on a paper, and something will happen and I didn't save it correctly and I will have to go through and re-write it and its irritating but its like I don't know I think I'd try to be more positive about than other people, it's a pain but at the same time it happens, there like, I look at it the same way as I do parking tickets, you are bound to get one eventually, I figure, it's sucks but its like at the point there not really anything you can do, you have to retype, I can't go back and get that file.

(Robert) So, I just saw it as a normal risk of doing business just like if I got a flat tire. You know, I could curse and jump up and down, but I would just get a new tire and move on.

Table 19: Description of Negative Digital Experience for all Participants

Negative Experience with Digital Information	Number of Participants
Technology Malfunction	12
User Error	7
System Incompatibility	6
Theft of Personal Information	1

Table 20: Emotions Associated with Negative Experience for all Participants

Emotions Associated with Negative Experience	Number of Participants
Frustration	9
Anger	7
Blame Self	4
Anxiety	2
Helplessness	2

Table 21: Change in Behavior/Beliefs Owing to Negative Experience for all Participants

Change in Behavior/Beliefs Owing to Negative Experience	Number of Participants
Save Content More Places	6
Change Behaviors for a Little While and then Return to Previous Behavior	5
Save Content More Frequently	4
Save or Copy Content More Carefully	4
Print Content	3
Save Content to Web Storage Space	2
Change the Way I think But not What I do	2
Purchased Anti-Virus Software	1
No longer uses a certain software	1

5.9 Discussion

The behaviors that contribute to the organization of personal information have limitations of scale. These behaviors are likely to inhibit access to personal information overtime given the paucity of metadata created to re-find content and the technical barriers to integrating browsing and/or searching of all personal content. Ultimately systems' design and individuals' organizational practices will need to be altered as the reliance on digital information is likely to continue or to increase. While preference for physical or digital formats is based on the context of the information need or use at hand, digital information formats are valued much more greatly than physical formats for meeting everyday information needs and uses.

6. Findings/Discussion: Part 3: Psycho-Social Implications

This chapter will explore the psycho-social implications of personal digital information preservation. Participants valued information in both the everyday and long term contexts because of its connection to other people. The values associated with those connections are discussed. Further, participants connected with other people to acquire knowledge related to information storage, management, and preservation.

6.1 Values Associated with Everyday and Long Term Information (N=25)

The importance of the values associated with everyday and long term information was calculated using the point system devised by Savolainen and Kari (2004) as described in Chapter 3. (See Tables 22-29.) Information valued in both the everyday and long term contexts connected the participants to other people. Information intended for long term use with emotional or affective value received the greatest value overall (145.5 points). In contrast, affective values associated with everyday information were much less (38 points). Items with emotional value included information items that represented family, friends, self, and sentimentality. Information intended for everyday use with a communicative value received the second highest value overall (87 points). In contrast, communicative value associated with long term information was much less (21.5 points). Information items with a communicative value include communication, accessing information, sharing information, and being informed. Documentation, creativity/aesthetics, and reference

as value were valued similarly for long term and everyday information. See Figure 5 for a comparison of points for all values.

The greatest single value was practicality of digital information in the everyday context (61 points). The next highest single values were emotional or sentimental (54.5 points) and then connection to family (50 points) in the long term use context. The emphasis on practicality speaks to digital information as a way of life in everyday information communication, creation and storage. Julie age 65 and Mark age 51, illustrate this point, even though they did not grow up using digital information but adopted its use over time.

(Julie) It's sort of transparent, because you get information from the tools you have or the tools you want to procure for that purpose, I have to use glasses to see, I think of those tools in the same, I don't think about oh I have to get glasses or I don't see well, it's just something I do. And that is the same way, I don't really think about it.

(Mark) I'm so used to having my computer with me that it's just, it's what I do, what I am.

(I) Okay.

(Mark) You know, I've been working with them for almost 30 years now… that when you talk about changing behavior – I know I would have information sitting in a book across the room, and I'll go to Google. (laughing) As opposed to getting up and move over and pick up the book and the information.

Affective values are attributed to selected personal information over time. This selected content becomes a conduit to memories that trigger emotional responses; whereas one could argue that information valued in the everyday context is valued for the content alone. Greg and Oscar illustrate this point below.

(Greg) Joyful…, I'd have to go back to pictures, that is the only thing I can think that is joyful and related to technology, it would be a picture that I took with my wife on the California coast.

(I) Why did you decide to take the picture?

(Greg) Because I saw her – the moment, I wanted to remember the moment.

Oscar, an anthropologist, worries that a dependence on digital information could be over encouraging the human brain's dependence on mental cues to form associations and remember life events.

(Oscar) Particular photos more in stuff that I've done with my daughter and watching her. I sort of had this fantasy when she was born that I would like have a book that I wrote, you know, "oh, and her first words were this", and none of that ever happened. So what I do have are photos of not necessarily here she is speaking her first word, but you know, like, oh yeah, I remember that period. I had totally forgotten when for years she wore a hat and never took her hat off. That is my repository of those memories. And I don't know if it's connected to the fact that I treat it as such, or just to the fact that I'm now 40, but I feel like my memory is getting significantly worse. And I suspect that it is connected to my reliance on digital technologies. I don't know what it is that my brain is being used for in place of that, but I know that, you know, certain societies where people don't, you know. I'm an anthropologist, so I've done tons of reading about other societies where there are people who can recite the entire history of their people word for word. There are people who the amount of knowledge that they hold; my brain seems to hold all kinds of archaic trivia and not much important knowledge – it's all stored digital.

Ideally individuals would attribute more metadata to information with emotional connection because mental associations fade over time and because the information item represents more than the content. By providing more metadata to information with affective value, individuals can store and preserve the narrative

context of their information over time. The following from Mark provides a good example where if extensive metadata were applied to photographs like the one he describes below then the family history associated with legacy photographs could be preserved as well as the image.

(Mark) My brother sent to me a digital scanned copy of my father's father's family. It's the only document we have that represents him as a child.

(I) That's amazing.

(Mark) So, it's a wonderful, wonderful piece.

(I) And the value of to you?

(Mark) It's very large because in my house, the whole hall, it's covered with family pictures that go back generations. And, it is all the branches of my family except for his branch.

(I) I see.

(Mark) So, I now have a piece that ties into his branch. Because, his mother died and his father left them at a very early age, and my great grandfather had to care for his siblings. He was 12 or 13.

(I) That happened a lot, right? The social systems were not in place to prevent that kind of thing from happening.

(Mark) Yes. On the other side of the family the mother died and the father dropped the three girls at a convent. So, they all became nuns.

The only values not attributed to both physical and digital information alike were monetary and historic. Participants specifically indicated only physical information items as having monetary and historic value. Perhaps not enough time has passed for digital information to be thought of in this way.

Table 22: Affective Values for all Participants

Affect	Everyday	Long Term
Emotional	13	54.5
Connection to Family	14	50
Connection to Self	10	30
Connection to Friends	1	11
Totals	38	145.5

Table 23: Documentation as Value for all Participants

Documentation	Everyday	Long Term
Legal	4	13
Medical	2	6
Evidence	4	3
Financial	28	4
Life Events	19	42
Totals	57	68

Table 24: Communication as Value for all Participants

Communication and Access	Everyday	Long Term
Communication	46	15.5
Access to Information	24	3
Sharing	12	3
Being Informed	5	0
Totals	87	21.5

Table 25: Reference as Value for all Participants

Reference/Support	Everyday	Long Term
Future Reference	18	20
Supports Work	20	11
Supports Creative Work	3	5
Supports School	3	0
Totals	44	36

Table 26:Uniqueness as Value for all Participants

Uniqueness	Everyday	Long Term
Irreplaceable	9	17
Difficult to Replace	8	10
Totals	17	27

Table 27: Function as Value for all Participants

Function	Everyday	Long Term
Saves Time	12	6
Practical	61	10
Use	7	0
Totals	80	16

Table 28: Creativity as Value for all Participants

Creativity	Everyday	Long Term
Creative/Aesthetic	21	27

Table 29: Historical and Monetary Value for all Participants

	Everyday	Long Term
Historical	0	15
Monetary	0	10
Totals	0	25

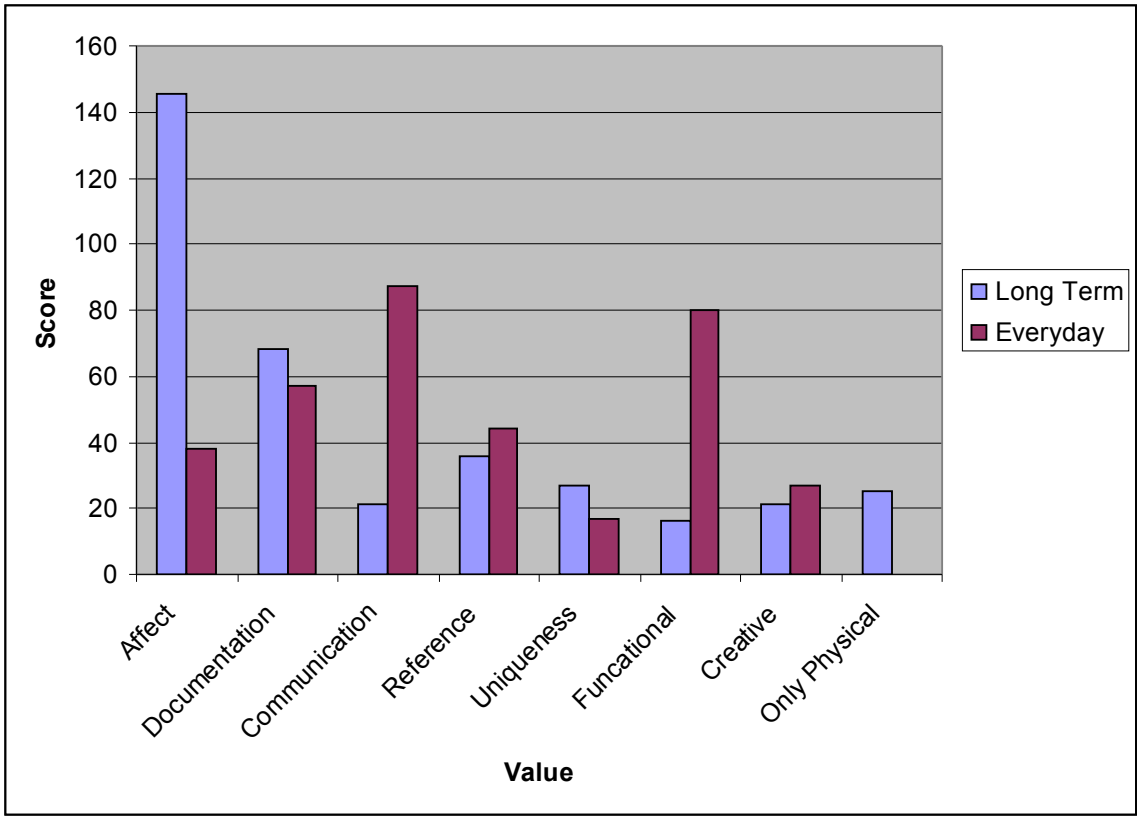


Figure 5: Comparison of Values for Long Term and Everyday Information

6.2 Positive Experience with digital information and the values associated with it N=26

In discussing a positive experience with digital information, participants described experiences that were valuable to them because of the ability to share, give or exchange digital files with others (N=11 or 42%) and because of their ability to be creative with technology (N=9 or 36%). (See Table 30 and Appendix K.) Six of the 11 positive experiences related to sharing involved photographs. Cathy describes the ability to share an experience abroad as it was happening, without the use of digital technology this would not have been possible.

(Cathy) Pictures of small town outside of Rome. My husband gave me a digital camera right before I left. I learned it on the fly. In order to tell somebody what I was doing I could send a picture to them which was enormously helpful. It was also incredibly useful to have the digital camera – because the place I was working in had restricted access. There were only five of us that were allowed in there. In order to show what somebody what the inside looked like, I could take a picture, so I was able to bride a gap.

(I) Would you share the photos via email?

(Cathy) Yes.

In the next example, Paula illustrates the ease with which sharing digital content takes place using email and social networking sites.

(Paula) I think my prom pictures, I remember all my friends had all their own separate digital cameras, we all sent all these pictures to everyone through email and I got to go through all the goofy ones and save them and delete the ones I didn't want.

(I) Some of these are your own and some are your friends?

(Paula) Yes

(I) You received them through email?

(Paula) Yes and we also used MySpace for some things

(I) Does your prom has a MySpace page?

(Paula) No we just put on them in our photos on MySpace, you'd just right click and save them.

(I) Why did you decide to save these?

(Paula) They are of my prom, it's just one of those things, it was just a very fun time. And the pictures are nice.

These two examples also illustrate email and social networking sites serving unintentionally as storage. The selection of content to share also indicates that the content is of some particular value.

This question drew a similar response from participants as the question asking them to describe a digital item of great value. (See Chapter 5, Table 16). Nine participants indicated that their positive experience with digital information involved their abilities to be creativity with technology. Uri provides a poetic description of this experience:

(I) Manipulating the photography and that kind of thing – what is that experience…what's the motivation there?

(Uri) I think you get…I'm going to use the wrong words but it's the same idea. It must be what Michelangelo or somebody who paints or somebody who does…and they have something and they say "hey, now that looks great." And it's something that they or I did. And, I'm not good with paint…I'm not good with art, but I've always liked that stuff. So I get, because of the computers, the ability to sort of create things that I like.

Lastly, three participants' positive experiences were tied the joy they experienced the first time they tried a new digital technology: using a digital camera, creating a playlist using iTunes, and chatting with friends. It is interesting to contemplate the existence of digital firsts. How many people have the transcript of their first chats or files of their first email ever sent or received? Who can remember the contexts of the events? A study of records documenting digital firsts would provide insights to the evolution and adoption of technologies currently in use.

Table 30: Values associated with Positive Digital Experience

Values	Number of Participants
Sharing	11
Creative/Aesthetics	9
Connection to Family	6
Documents Life Events	6
Supports Work	4
Connection to Friends	3
Enjoyment of Technology	3
Practicality of Digital Information	1
Communication	1

6.3 Knowledge of Digital Storage and Formats (N=22)

Participant knowledge of digital storage and formats is discussed in this chapter because of the importance of social connections in acquiring knowledge related to digital technology which is discussed in sections 6.4 and 6.5 of this chapter. Also, the question related to the participants' positive experience with digital information is used illustrate knowledge of file formats, which is discussed on section 6.2.

In general, participants lacked specific knowledge regarding which storage devices/locations or file formats were considered best for providing long term access to digital content. Half the participants selected redundant storage space based on ease of use (N=11 or 50%). (See Table 31.)While participants were mainly motivated to save information for future reference (N=15 or 68%), when selecting redundant storage reliability of the storage device was not considered.

Below, Nora and Zoey discuss why they selected their redundant storage devices. Their comments indicate that there is some cost associated with storing digital information. Through her comments, Zoey describes a common path that many consumers have taken as more flexible and easier to use storage devices were developed.

(Nora) Cheapest.

(I) Cheapest. Okay.

(Nora) You know, it's like, we have a three year old Mac, and so one of the things is it needs more memory to hold all of this stuff and we would have to pay like $500 to get an external hard drive to get more memory and we're going to have to buy a new computer soon, so I don't want to do that. So, we chose the disks just because it's the quickest and, you know, just in case.

(Zoey) The flash drive for the convenience of it, I suppose. Before they came around I thought zip drives were like the wave of the future.

(I) Oh, so did everybody else.

(Zoey) So, I had the zip drive. I was really into that. But then, otherwise I would sometimes back up to CD. But, I felt like that wasn't very… because you can't…you'd have to erase the whole thing to rewrite it. So, flash drives are kind of like floppies and they are zip drives, but they hold more. So, pretty useful.

(I) Okay. Why did you pick the particular one you bought?

(Zoey) I guess for size and probably for price.

Table 31: Reasons for Selecting Redundant Storage Space for all Participants

Reason for Selecting Redundant Storage Space	**Number of Participants**
Easy to Use	11
Portable	7
Storage Capacity	7
Access Speed	5
Price	5
Integrates with Other Technology	4
Recommended by Someone Else	4
Offsite Storage	3
Privacy	2
Trial and Error	2
Unintentional	2

Further, most of the participants possessed limited knowledge of the file formats they used and of the importance of considering file formats when preserving digital information.

Whenever participants mentioned a particular digital item of value, they were asked questions regarding the file format of that particular item. For example, when participants discussed a positive experience with digital information they were then asked questions about the format of the item. (See Table 32.) In ten instances, the participants did not know the file format for the item being discussed, and in 19 instances the participants did not know if the format was recommended for long term access. (See Table 33.) Only two participants were knowledgeable regarding all the formats they used and the reasons why they were or were not good for preservation. (See Table 34.)

In many cases, participants depended on the device to determine the format and gave no thought to this aspect of managing digital content. In 21 out of the 26, or 84%, positive digital experiences discussed, the device or software being used to create the item determined the format. Digital cameras and Microsoft software were the most common devices/software applications used. For all 26 participants, the devices or the software used by them determined at least some of the formats of information items they create. (See Tables 33-34.) The discussion with Irene below demonstrates a lack of an awareness of the need to consider certain technical aspects in order to preserve her personal information prior to her interview.

(Irene) Whatever dot something.

(I) You don't know what type of file it is?

(Irene) No some kind of movie making thing.

(I) And do you have that backed up anywhere?

(Irene) No, I'm going to go back up everything now.

(I) Yeah, that's the effect I have on people.

(I) Your not sure of the format, you don't have it backed up and you're not sure about the quality of the CD?

(Irene) Yes, that's it.

Seven participants expressed that their lack of attention to formats prevent them from accessing their content "forever" as in this example from Zoey.

(I) Okay, and how long do you plan on keeping these photos?

(Zoey) I guess forever.

(I) Okay, and how do you care for them?

(Zoey) I don't.

(I) Do you know what format they are in?

(Zoey) They are in JPEG.

(I) Okay. Do you think that's a good format for long term preservation?

(Zoey) Maybe not. I mean, I guess I could change. Something newer could come along, better way to save photos. It could become obsolete.

Table 32: Knowledge of Digital Formats of Items discussed during Positive Experience with Digital Information by Individual Participant

	Content Type	Format	Knowledgeable about Formats Good for Preservation	Device Default
Ann	Photos	TIFF and JPEG	Yes	No
Bill	Photos	JPEG	No	Yes
Cathy	Photos	Unknown	No	Yes
David	Emails	Unknown	No	Yes
Eleanor	Photos	JPEG	No	Yes
Frank	Web pages	HTML and XML	Yes	No
Greg	Photos	JPEG	No	Yes
Holly	Text Messages	Unknown	No	Yes
Irene	Digital Video	Unknown	No	Yes
Julie	Emails	Word Doc	No	No
Kelly	Photo Story	Format from Microsoft Software	No	Yes
Lisa	Document	Unknown	No	Yes
Mark	Photos	JPEG	No	Yes
Nora	Image	JPEG	No	N/A
Oscar	Music	AAC and MP3	Yes	Yes
Paula	Photos	JPEG	No	Yes
Quincy	Web pages	HTML	No	Yes
Robert	Chat	Unknown	No	Yes
Sarah	Photos	JPEG	No	Yes
Tina	Image	Format from Microsoft Software	Yes	Yes
Uri	Photos	TIFF or Raw	Yes	No
Victoria	Document	Word Doc	No	Yes
Wendy	Photos	Unknown	No	Yes
Xavier	Document	Word Doc	Yes	Yes
Yvonne	Photos	Unknown	No	Yes
Zoey	Document	Word Doc	No	Yes

Table 33: Specific Knowledge related to Positive Experience

Knowledge Related to Positive Digital Experience	Number of Participants
No Knowledge of File Format Type	10
Device Determined the File Format	21
No Knowledge if Particular Format is Good for Long Term Use	19

While most of the participants (N=17 or 65%) completely lacked knowledge regarding digital formats and preservation, some participants (N=7 or 28%) were knowledgeable of only certain formats and two participants were very knowledgeable regarding the formats they used. This discussion from Oscar demonstrates the depth of his knowledge and the need for formats that can be accessed independent of device or software.

(Oscar) I wonder. You know, I have thus far, for the most part, been able to preserve my digital information but translating it into other standards. But for example, this is outside of what we're talking about right now, but I know that in the last operating system Apple had had all of the emails in some kind of public format. And that has changed. You know, they've… that was in theory it was a response to people complaining about their proprietary format that Apple Mail used. But then in 10.5 they went back to a proprietary format and – I didn't even notice this – but somebody pointed out that he bought the new system, he – this is some blog I read – he popped it in, he opened mail for the first time and it said "converting your information." But he no longer has all of those emails that he had worked so hard to get into a public format. Now they've just been translated into a proprietary format. Now, .aac, I don't know if that's proprietary to Apple, to be honest. I don't think it is. I mean it's Apple _____ or whatever…no, that's something else. But, its common enough and I have the technology to transfer it to .mp3. I feel like .mp3 will be with us for a while and it's so common that whenever it's not with us I feel like there will be huge numbers of people who publish translation programs that get it into whatever the new format is. But I do suspect that ultimately I'll lose…you know, that's why I keep my CD's. They're a much higher quality than either of those

Table 34: General Knowledge related to Formats Good for Long Term Use

Preservation Knowledge Related to Formats Used	Number of Participants
No Knowledge of formats Used	17
Knowledgeable about some but not all formats Used	7
Knowledgeable about all Formats Used	2
Awareness that lack of knowledge could be a problem	7
Allows devices to determine Formats at least some of the Time	26

6.4 Knowledge Acquisition relevant to the Management of Personal Digital Information (N=26)

Participants turned most frequently to other people for information on information storage, management, or preservation: family and friends (N=8 or 31%), retail professionals (N=5 or 19%), community of practice (N=5 or 19%), and co-workers (N=4 or 15%). The only other source consulted was the internet (N=8 or 31%). (See Table 35.) Three examples are provided in the excerpts below. Paula turns to a skilled family member; Oscar uses the internet and then consults with a retail professional; and Ann uses the internet and when that doesn't work she turns to a community of practice.

(Paula) when it comes to my pc at my house, I always ask my brother, it was being weird a few months a go, it was going really slowly and he fixed it. He's actually the big computer guy at a national not-for-profit. He knows everything.

(Oscar) Well, the obvious example is where am I going to put my music library? And that led to management and storage, actually. I mean, that was the most recent time I was trying to figure out, okay, so now I've outgrown my iPod and it doesn't fit. Who did I turn to? I think just turned to the web. And I think I stopped by the Apple store when

I was at Trader Joes and asked some questions before I went back to buy.

(Ann) I enjoy finding the solutions myself but I'll have to tell you when I stop enjoying that I have friends I call who know, who know they know Final Cut Pro. They use it everyday. Why would I not call them and ask them? There are so many sites; if you can't find the answer then you are truly lost. As for as MS Office, any application they have, I'll have to say they are pretty good about their online and offline help for what I need. I'm not creating databases everyday.

Most participants would prefer to learn about the technology that they use with the aid of a personal instructor (N=14 or 54%) and secondary to that participants preferred teaching themselves (N=6 or 24%) typically using the internet. (See Table 36.) Kelly, Oscar, and Quincy discuss their preference learning while interacting with others and the technology.

(Kelly) I learn better from being in a class or having someone explain to me and show me how to do something. If I have to read something like directions on how to do something online, but on the computer, there are so many different components, I get confused and it's irritating. If someone can sit down and explain it to me and show me it's easier for me to recall and to understand.

(Oscar) I don't know. I mean, I'm most successful at just Google-ing stuff and figuring it out and finding online discussion boards where people have hashed their own way through their problems and trying to find the one that's geared towards my level of knowledge. I guess in an ideal world where money wasn't a factor, there is some of that that I would hire a tutor to do, you know. Like if my friend Spud wasn't busy running his own business I'd say "Spud, how do we do this?" and he would show me. But I never, it never makes sense unless I'm doing it. I can't learn that kind of knowledge from a book. I'd have to play with it and say oh, well that didn't work, well okay let's see, apparently I'm still getting "there's a network address translation error" up here so I've got to open a port on my router.

(Quincy) The best way for me to learn it? Practice. Actually doing it. I'm not asking for more opportunities to run into problems so I can

> troubleshoot. But, I think that hands on experience is going to stay there. It's not…I won't just jot it down, push a couple of buttons, and forget about it. I think doing it over and over again, you know, so maybe a class, some type of hands on class, where we're doing projects – that kind of thing. I don't do that well when they have the video tutorials online. I'm having some trouble right now, as we speak, trying to upload some stuff in MP3 format and it kills me. But, I wouldn't want to sit in front of a computer and learn about technology, using technology in that way. I'd rather have someone saying "Okay, here's what you do" and be able to touch it and manipulate it and, you know, really get that hands on experience.

Table 35: Information Source Used for Information Management and Storage Questions for all Participants

Information Source	Number of Participants
Family & Friends	8
Internet	8
Retail Professional	5
Community of Practice	5
Co-workers	4

Table 36: Preferred Method of Learning about Digital Technology for all Participants

Learning Styles	Number of Participants
Personal Instructor	14
Self	6
Classroom Setting	5
Hands On	5
Reading	2

6.5 Discussion

Information valued in both the everyday context and long term context connected the participants to other people. In the everyday context, the value was associated with the communication of information to other people. In the long term

context, the value was connected with emotions associated with friends and family through the documentation of life events. Physical information items of great value typically point outward to others and were object or product oriented. Whereas, digital information items of great value typically point inward to self and were connected to the process of creation and one's own talents with technology. Values converged around information desired for preservation regardless of format. The values include: sentimentality, connection to family, connection to self, financial, legal, and creativity/aesthetics. The only two values that were not associated with digital information deserved for preservation are monetary and historical.

The participants' knowledge of preservation practices related to storage media and file format selection was limited. Knowledge regarding information management or preservation was mostly obtained from other people: family, friends, retailer employees, co-workers. The only other form of information consulted was the internet. One-on-one instruction was the most preferred way to learn about anything new related to technology.

7. Conclusion and Implications for Future Research

7.1 Summary of Findings

The following three sections (7.1 – 7.3) provide a summary of the findings presented in chapters four through six. These sections also serve to detail the components that influence the conceptual model of personal digital preservation practices featured in section 7.4. The technical, cognitive and affective, and psycho-social components of the model provide an understanding of the preservation practices of public library users. The implications of this understanding for future research will be discussed later in this chapter.

The components emerged from the data. They are based on the participants' behaviors and beliefs associated with the preservation of their digital information given the constraints imposed by systems and access to consumer level information related to digital preservation practices.

7.2 Technical Components

In response to research question one, regarding the what content types individual participants saved and why it was found that content formats are mostly decided by digital devices and or/Microsoft Office software rather than personal preference or knowledge of preservation practices. Content is most likely to be stored where it is created which is typically on the stable storage media found in desk tops, lap tops, and external hard drives. But this is not always the case as client side device driven access limits access and storage capacity.

Digital devices place limitations on individuals creating and storing digital content, like digital video and music, which require large storage capacity. The need for flexible storage and easier resolution to interoperability issues may encourage a move towards web based creation and storage practices. Also, content is unintentionally being stored on the web as a result of social interactions through email and social networking sites like Facebook. While this further distributes personal content, the sharing of information may indicate value, may be easier to find again given the social conversation, context and networks associated with the information.

In response to research question two, technology design and use informed the steps individual participants could take to preserve their own content. Content is distributed across many devices and web locations with no technical means of unifying access and identifying comprehensively all personal digital information. In time, this disorganization will inhibit access as cognitive associations to location of personal information will fade, especially given the practice of using the current context to classify and organize personal information.

Making duplicate copies of personal content is the principle preservation strategy employed. This is a result of individuals losing content due to technology malfunctioning or inoperability issues related to technology upgrades. Digital devices are viewed as generally stable but also as bound to fail or disappoint eventually. As a result, email is used intentionally as a storage location to store duplicate copies of content and for ease of access.

7.3 Cognitive and Affective Components

In response to question two, regarding the steps individuals are taking to preserve their own digital content, it was found that Personal unifying taxonomies are not created or maintained. Further, no technology exists to unify access to content stored on multiple devices in and multiple locations. Browsing to re-find personal content is the norm due to the limitations imposed by digital devices on metadata creation and search functionality.

Preference for and use of certain information formats has the potential to inform preservation decisions. Preference for physical or digital formats is based on the context of the information need or use at hand. However, digital information formats are valued much more greatly than physical formats for meeting everyday information needs and uses. Information use is an indicator of value for everyday information to a great extent and to a lesser extent information valued for long term use.

In response to question three, regarding how participants could improve their preservation practices, it was found that file and folder classification is inspired by the current context of the information use. When imposed, organization strategies are informal and idiosyncratic. Organization relies heavily on broadly classified folders to order and define content topicality. The use of metadata is limited and typically involves the creation of current context inspired file names. Photographs are often only identified at the folder level, leaving this type of content particularly vulnerable to loss through the loss of mental associations. The previously mentioned behaviors have limitations of scale. These practices are likely to inhibit access to personal

information over time, given the paucity of metadata created to re-find content and the technology barriers to integrating browsing and/or searching of all personal content.

7.4 Psycho-Social Components

In response to research question one, what participants saved and why, the following was found. Most positive experiences with digital information result from sharing or exchanging digital content with family and friends. In most cases, general categories of information are indicated as valuable rather than specific items. It is widely supported that long term memory privileges the organization of information through general associations held together through established mental models or schemas.

Physical information items of great value typically point outward to others and are object or product oriented. Whereas, digital information items of great value typically point inward to self and are connected to the process of creation and one's own talents with technology. Values converge around information desired for preservation regardless of format. The values include: sentimentality, connection to family, connection to self, legal, and creativity/aesthetics. The only two values not associated with digital information desired for preservation are monetary and historical. Information valued in both the everyday context and long term context connects individuals to other people. In the everyday context, the value is associated with the communication of information to other people. In the long term context, the

value is connecting with emotions associated with friends and family through the documentation of life events.

With regards to research question three, how could individuals improve their knowledge and what would motivate them to improve, the following was found. Knowledge of preservation practices related to storage media and file format selection are limited. Knowledge regarding information management or preservation is mostly obtained from other people: family, friends, retailer employees, co-workers. The only other form of information consulted is the internet. One-on-one instruction is the most preferred way to learn about anything new related to technology.

7.5 Conceptual Model of Digital Preservation Practices

A conceptual model based on the components identified above illustrates of the process of digitally representing life experiences and the preserving of those representations through documentation. (See Figure 6). Based on the data collected for this study, an individual's cognitive, affective and social reactions and interactions inform her life experiences. In turn, these factors combine to influence the creation, acquisition, organization, and ultimately the preservation of personal information.

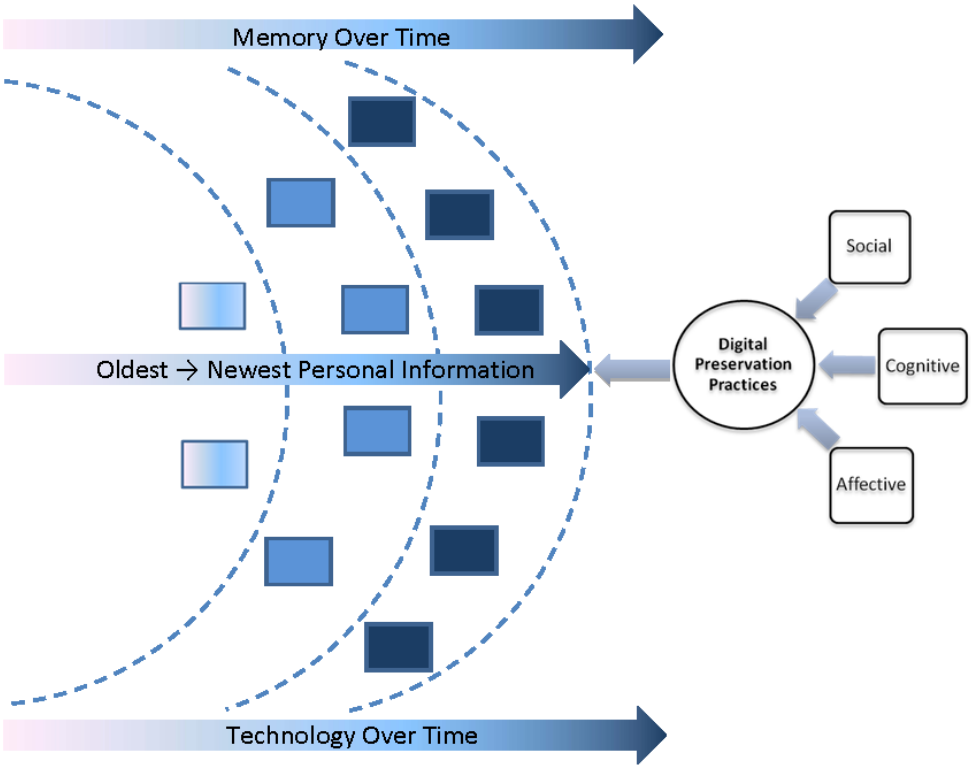


Figure 6: A Psycho-Social-Technical Model of Personal Digital Preservation Practices

Technological affordances facilitate the creation, storage, and accession of digital information. However given the current design of personal digital devices, affordances have the potential to become constraints over time. Identifying and re-finding personal digital information becomes more difficult as more content is stored and distributed among various web locations and devices. With the passage of time, an individual's cognitive associations to specific life experiences fade. Documentation of new experiences begins to out pace older documentation. As mentioned previously, the participants' relied on informal organizational practices

and on the current context for naming files. These practices offer limited support for cueing human memory to recalling content of specific value.

Further, technology continues to evolve, leaving older documentation/digital content susceptible to obsolescence or degradation. Fortunately, as discussed in Chapter three, current digital preservation research is attempting to address the access problems presented technological obsolescence and media and file degradation. Therefore it is important to incorporate of the study of affective, cognitive, and social factors with technological factors related to digital preservation for a viable and comprehensive solution to the problem.

7. 6 Implications for Systems Design

As personal digital devices were not initially designed as conduits to information retrieval or archival access, there is much potential for improving long term access to personal digital information through improvements to the design of personal digital devices. The problem of scale associated with re-finding personal information has the potential to be addressed through changes to systems' design. Further research needs to explore the development of tools to support individuals in the organization and preservation of personal information.

Based on the model and its components, the following areas could be improved through better system design: metadata attribution, content unification, and automatic identification of valuable information. In most cases, general categories of information were indicated as valuable rather than specific items. Given this finding, research regarding the types of metadata needed to record and preserve the context of

the personal/family narrative along with personal digital content is needed. Also, research is needed on how best to design systems to encourage and facilitate the application of metadata.

Content is distributed across many devices and web locations with no technical means of unifying access and identifying comprehensively all personal digital information. In time, this disorganization will inhibit access as cognitive associations to location of personal information will fade over time, especially given the practice of using the current context to classify and organize personal information. Research focused on the design of tools that would unify content, content metadata and content location is needed to reduce the fragmentation of personal information.

The values associated with digital information are likely to evolve over time. Continued study of cognitive and affective interactions with digital information needs to take place, especially long term studies. In conjunction with this research, the development of tools and systems that can help individuals identify digital information of value needs to occur. If a tool could automatically detect value through, for example, frequency of use, number of redundant copies, or the application of key metadata, then a means of mining personal digital information of value would be possible.

Given that information shared is information selected, social networking sites have the potential to aid in the automatic organization and identification of personal content of value. Personal content shared through email exchanges could also be a means of identifying content of value. To determine the extent of the significance of

social networking sites and email exchanges in identifying content of value would require future research.

7.7 Implications for Storage Selection

Not one participant mentioned losing content do to a web storage malfunction yet most every participant had a digital loss due to devices malfunctioning. Further research needs to explore if in fact web based storage is more reliable and if it is, what would encourage individuals to store content on the web over digital devices? Further, the examination of individual preference related to storing content would help the design of personal storage devices/locations. Is the preference a matter of client-side software privileging client-side storage over a more flexible web storage or do individuals feel that networked storage inherently comprises personal privacy?

A study of access problems encountered with personal digital content stored on a variety of device types and web locations would help to identify problems that individuals are likely to face over time and to illuminate content potentially at risk.

The National Diet Library of Japan study on packaged digital publications acquired from 1980 to 2002, provides some indication of the problems individuals are likely to have accessing their stores of digital information. Of the 200 items in their study, 70% presented access problems. In 50% of the cases, problems with operating system incompatibility prevented access to the packaged digital item. In 30% of the cases, it was a problem specific to the application that plays back the media. The problem was specifically reported as the latest version of Adobe Reader being incompatible with plug-ins specific to older media. Only 12% was specific to the media type breaking down. The study found that the longer the item had been in the

library's collection the greater the chance that the medium itself would prevent use (National Diet Library, 2008).

7.8 Implications for Public Libraries

Considering the participants' preservation practices and learning preferences, research is needed to explore how public libraries can educate their constituents to better care for their personal digital content. Participants indicated that they preferred to learn through hands-on, context driven learning situations aided by a personal instructor. Japzon (2008) has suggested that public libraries and other local memory organizations work with their constituents to create community repositories for the preservation of personal information, which also contributes to social and cultural histories.

From the museum science community, Russo and Watkins (2007) describe "community co-creation" as cultural institutions and communities working together to create digital content, each benefiting and learning from the other's expertise and experience. Library and information professionals working through their organizations could provide a technical infrastructure and contribute technical expertise in collection development and maintenance. Community members could provide the knowledge that supports content development and contribute personal digital information and physical artifacts from their own collections. Future research would be needed to determine the efficacy of community repositories for teaching and learning.

Public libraries have the potential to provide consumer level information on digital preservation accessible to all of their constituents. However, certain populations may require further study to adequately address their needs. Given the finding that men were more likely to systematically back up their own content as well as the content of their spouses, future research on the influence of gender on the acquisition of technology related knowledge could prove helpful in identifying cultural biases that might exist to encourage differences.

Also, the study required that participants owned a personal computer. This requirement excluded individuals who do not own a personal computer and rely on the public library to meet their personal computing needs. Research is needed to identify the digital preservation challenges inherent to those affected by the digital divide.

Lastly, the comparison of all digital collection efforts (institutional and personal) and the role that digital preservation plays in each of these collections could provide greater insight regarding collection development, maintenance and preservation of digital information than can be gained than looking at these collections separately.

7.9 Major Contributions of Research

This research provides an understanding of public library users' approach to digital preservation. In doing so, this research fills a gap in the digital preservation research which to date has focused on diplomatic, scholarly, and corporate information.

Further, this research explored the affective, cognitive, social reactions and interactions with personal digital information and the technology that supports it. It provides a holistic model for studying other groups and their digital preservation practices. Also, the study of the motivations for content acquisition and storage source selection along with the values associated with personal digital information extends information behavior research beyond information seeking, searching, and use. Lastly, the research indentified the gaps in the knowledge possessed by individuals regarding digital preservation practices.

7.10 Conclusion

This study found that participants overwhelmingly valued digital formats, rather than physical formats, for everyday information contexts, and valued physical formats only slightly more for long term information contexts. The growing dependence on digital information for communication and documentation may be changing the way human memory works. Research regarding the relationship between the reliance on digital information for memory recall and the reliance on human memory alone for recall needs to be explored in relation to personal histories as well as social histories.

The diminishing role of physical and mental "touch" in information acquisition, organization and storage has implications for long term access. If the associations are formed through "touch" are absent or diminished, then the processes that engage the working and long term memory through these associations are likely to be absent or diminished as well. Research needs be conducted to confirm the effect

of digital information creation and use on the formation of cognitive associations. If in fact digital memory is replacing cognitive memories, even to a certain extent, the challenges of digital preservation and long term access to information will need to become a greater research issue for the library and information science community than it is now.

The growing dependency on and preference for digital formats has implications for the production and storage of information sources. How much longer will publishers and libraries have the economic means to produce and provide the same content in a variety of formats? For example, is there a need for publishers continue to produce physical copies, mp3 files, audio discs, audio tapes, HTML files, PDF files, etc – all of the same content? Likewise, will libraries be able to justify the purchase and storage of the same content numerous times over, just in different formats? Better technologies will likely win out. As the use of digital information for communication and documentation becomes the norm rather than a choice, if it hasn't already, reliable, affordable and trusted stores of digital information will need to be developed to record personal and social histories. The cost of digital information production and access will continue to decrease to the point where the major costs will be related to preservation. This is where the role of public libraries may shift primarily from information access and life long learning or evolve to include the provision of preservation as a public good.

Libraries have kept apace of the technological changes to information formats in order to meet the expectations of their constituents. Most libraries provide access to digital information in a variety of formats and delivery options. But at some point

in the future, providing access to digital information may no longer be primarily the role of libraries, especially if Google Scholar and Books are harbingers of the types of information sources to come. Also, as the production and publication of digital information increasingly merge, the need for information intermediaries to select, acquire and provide access to information is decreasing (Janes, 2003).

Libraries evolved from their primary role of ownership of information to one of primarily providing access. Libraries may return once again to primarily being concerned with information ownership because of the need to provide digital preservation as a public good: long term access to our collective culture heritage is not a commodity for sale. The responsibility and cost of digital preservation is too great for anyone person, institution, company or market place to bear.

For libraries and other memory organizations to create a sustainable infrastructure for digital preservation they must collaborate with each other to create interoperable systems, and develop standards for creation and storage (Bradley, 2007; Mason, 2007). Ideally, any preservation infrastructure would evolve to include public library users and the preservation of personal digital information. Otherwise, who will individuals collaborate with to preserve their own personal information, how will individuals know what the best practices for storage and file formats are, and in what archival infrastructure will such valuable personal digital information reside?

Public libraries could work with their constituents to support the preservation of personal information, which also contributes to social and cultural histories. Collection creation, maintenance, and preservation (physical or digital), are knowledge intensive processes. Public libraries could provide community members

with trusted sources to seek information related to digital information management and preservation. Libraries and archives have a tradition of helping individuals care for personal tangible information such as letters and photographs.

Generally, public libraries have connected with users by providing access to computers and networked proprietary resources for personal use, but have not succeeded nearly as well at connecting with individuals in their personal computing spaces. Further research is needed on the role of community members in public library digital collections, and on the connections between the expertise of information professionals and the knowledge needed to manage personal information collections over the long term. In this way, librarians and LIS researchers can become a driving force behind helping the public to preserve and organize their personal information collections.

References

Abrams, D. (1997). Human factors of personal web information spaces. Technical report, *Knowledge Media Design Institute Technical Report 1*, University of Toronto.

Abrams, D., Baecker, R.M., and Chignell, M. (1998). Information archiving with bookmarks: Personal web space construction and organization. *Human Factors in Computing Systems: Proceedings of the CHI '98 Conference,* (p. 41-48)*,* New York: ACM.

Abrams, S. L. (2005). Establishing a Global Digital Format Registry. *Library Trends, 54*(1), 125-143.

Agosto, D. E. & Hughes-Hassell, S. (2005). People, places, and questions: An investigation of the everyday life information-seeking behaviors of urban young adults. *LISR, 27*, 141-163.

Alemneh, D. G., Hastings, S. K., & Hartman, C. N. (2002). A metadata approach to Preservation of digital resources: The University of North Texas Libraries' experience. *firstmonday*, (7)8. Retrieved from http://www.firstmonday.org/issues/issue7_8/alemneh/index.html on November 13, 2006.

Allen, D. & Wilson, T. (2003). Information overload: Context and causes. *The New Review of Information Research: Studies of information seeking in context. Proceedings of ISIC, 2002*, 4, 31-44.

Anderson, N., Hodge, G. & Japzon, A. (2007). Harnessing NASA Goddard's grey literature: The power of a repository framework. *The Grey Journal, 3*(3), 154-158.

Balter, O. (1997). Strategies for organizing email messages. In *Proceedings of HCI '97Conference on People and Computers XII.* Springer. Pp. 21-28.

Barksdale, J. & Berman, F. (2007). Saving our digital heritage. Washington Post. May 16, 2007. p. A15.

Barreau, D. (1995). Context as factor in personal information management systems. *JASIS, 4*6(5), 327-339.

Barreau, D. & Nardi, B. (1995). Finding and reminding: file organization from the desktop. *ACM SIGCHI Bulletin, 27*(3), 39-43.

Barry, C. (1994) User-defined relevance criteria: An exploratory study. *JASIS 45*, 149-159.

Barry, C. & Schamber, L. (1998). Users' criteria for relevance evaluation: A cross-situational comparison. *Information Processing & Management, 34*(2-3), 219-236.

Bartlett, F.C. (1932). *Remembering: An Experimental and Social Study*. Cambridge: Cambridge University Press.

Bartlett, F.C. (1958). *Thinking*. New York: Basic Books.

Beagrie, N. (2005, June). Plenty of room at the bottom? Personal digital libraries and Collections. *D-Lib Magazine 11*(6). Retrieved from http://www.dlib.org/dlib/june05/beagrie/06beagrie.html on February 3, 2007.

Belkin, N.J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, 133-143.

Bergman, O., Beyth-Marom, R. & Nachmias, R. (2003). The user-subjective approach to personal information management systems. *JASIS&T, 54*(9), 872-878.

Besser, H. (2000). Digital longevity. In *Handbook for Digital Projects: a Management Tool for Preservation and Access*. Andover, MA: Northeast Document Conversion Center.

Bellotti, V. & Smith, I. (2000). Informing the design of an information management System with iterative fieldwork. In D. Boyarski & W. Kellogg (Eds.), *Proceedings of the DIS 2000 conference on Designing interactive systems: Processes, practices, methods, and techniques* (p.227-237). New York: ACM Press.

Boardman, R. & Sasse, M. A. (2004). "Stuff goes into the computer and doesn't come Out" a cross-tool study of personal information management. *CHI 2004*. April 24-29, 2004. Vienna, Austria.

Borgman, C. L. (2003a). *From Gutenberg to the global information infrastructure: Access to information in the networked world.* Boston: MIT Press.

Borgman, C. L. (2003b). Personal digital libraries: Creating individuals spaces for innovation. NSF Workshop on Post-Digital Libraries Initiative Directions. Workshop Paper. Retrieved from http://www.sis.pitt.edu/~dlwkshop/paper_borgman.pdf on January 31, 2007.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Routledge, London.

Bradley, K. (2007). Defining digital sustainability. *Library Trends, 56*(1), 148-163.

Branin, J. (2005). Institutional repositories. In *Encyclopedia of Library and Information Science*. Ed. Miriam Drake. Pp. 237-248.

Bruce, H. (2005, April). Personal, anticipated information need. *Information Research, 10*(3). Retrieved from http://informationr.net/ir/10-3/paper232.html on February 10, 2007.

Bruce, H., Jones, W., & Dumais, S. (October, 2004). Information behavior that keeps found things found. *Information Research, 10*(1). Retrieved from http://informationr.net/ir/10-1/paper207.html on February 10, 2007.

Buckland, M. (1991). Information as thing. *JASIS, 42(5),* 351-360.

Bush, V. (1945). As we may think. *Atlantic Monthly176*(1): 101-108.

Caplan, P. and Guenther, R. (2005). Practical preservation: The PREMIS experience. *Library Trends, 54*(1), 111-124.

Carr, N. (2008). *The Big Switch: Rewiring the World from Edison to Google.* New York: W. W. Norton & Company.

Carroll, J. (1982). Creating names for personal files in an interactive computing environment. *Journal of Man-Machine Studies, 16,* 405-438.

Case, D. O. (1991). Conceptual organization and retrieval of text by historians: The role of memory and metaphor. *JASIST, 42*(9), 672-689.

Chakraborty, J. & Bosman, M. 2005. Measuring the digital divide in the United States: Race, income, and personal computer ownership. *The Professional Geographer, 57*(3), 395-410.

Chen, Su-Shing. (2001, March). The paradox of digital preservation. *Computer*, 24-28.

Choi, Y. & Rasmussen, E. (2002). Users' relevance criteria in image retrieval in American history. *Information Processing & Management, 38*(5), 695-726.

Cunningham, A. (1999). Waiting for the ghost train: Strategies for managing electronic personal records before it is too late. *Archival Issues: Journal of the Midwest Archives Conference*, *24* (1), 55-64.

Cunningham, S. J. & Masoodian, M. (2007). Identifying personal photo digital library features. *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital*

Libraries (pp. 400-401). New York, NY: Association for Computing Machinery

Cutrell, E., Dumais, S. T., & Teevan, J. (2006, January). Searching to eliminate personal information management. *Communications of the ACM, 49*(1), 58 – 64.

Damasio (2006). Descartes Error. New York: Vintage.

Davis, S. E. (2008, Spring/Summer). Electronic planning in "collecting" repositories. *The American Archivist, 71,* 167-189.

Day, M. (1997). Extending metadata for digital preservation. *Ariadne, The Web Version.* Retrieved from http://www.rlg.org/preserv/joint/day.html on April 9, 2007.

Day, M. (1998). Issues and approaches to preservation metadata. *Proceedings of the Joint RLG and NPO Conference: Guidelines for Digital Imaging.* Retrieved from http://www.rlg.org/preserv/joint/day.html on April 9, 2007.

Day, M. (2004). Metadata applications and management. In *International Yearbook of Library and Information Management*. Eds. G. E. Gorman and Daniel G. Dorner. London: Facet Publishing, pp. 253-273.

Dervin, B. (1992) From the mind's eye of the user: The sense-making qualitative quantitative methodology. In J. D. Glazier & R. R. Powell (Eds.), *Qualitative Research in Information Management* (pp. 37-50). Englewood, CO: Libraries Unlimited, Inc.

Dix, A. & Marshall, J. (2003). At the right time: when to sort web history and bookmarks. In *Proceedings of Human Computer Interaction International 2003*.

Ducheneaut, N. & Bellotti, V. (2001). E-mail as habit: an exploration of embedded Personal information management. *Interactions, 8*(5), 30-38.

Elsweiler, D., Ruthven, I., & Jones, C. (2007). Towards memory supporting personal information management tools. *JASIS&T, 58*(7), 924-946.

Eppler, M. & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society, 30*, 325-344.

Holsti, O. R. (1969). *Content Analysis for the Social Sciences and the Humanities.* Reading, MA: Addison-Wesley.

Friedman, B., Kahn, P. H., Jr., & Borning, A. (2006). Value sensitive design and information systems. In Zhang, P., & Galletta, D. (Eds.), *Human-computer interaction and management information systems* (pp. 348-372). Armonk, NY: Sharpe.

Gasson, S. (2004). Rigor in grounded theory research: An interpretive perspective on Generating theory from qualitative field studies. In *The Handbook of Information Systems Research*. Eds. Michael E. Whitman & Amy B. Woszczynski.
Hershey, PA: Idea Group, Inc. Pp 79-101.

Gemmell, J., Bell, G. & Lueder, R. (2006). MyLifeBits: Personal database of everything. *Communications of the ACM, 49*(1), 89-95.

Glaser, B. B & Strauss, A. (1967). The Constant Comparative Method of Qualitative Analysis. In *The discovery of grounded theory: Strategies for qualitative Research* (pp. 101-115). Piscataway, NJ: Aldine Transaction.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report, 8*(4), 597-607.

Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ.* New York: Bantam Books.

Gonzales, R. & Woods, R. (2007). *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ: Prentice Hall.

Gorman, G. & Clayton, P. (2005). *Qualitative Research for the Information Professional.*2nd Edition. London: Facet Publishing.

Gottlieb, L. & Dilevko, J. (2001). User preferences in the classification of electronic Bookmarks: Implications for a shared system. *JASIS&T, 52*(7), 517-535.

Hafner, K. (2007). Analog memories in a digital world. *New York Times*. March 10, 2007.

Hedstrom, M. (1997). Digital preservation: A time bomb for digital libraries. *Computers and the Humanities, (33)*3, 189-202.

Henderson, S. (2005). Genre, task, topic, and time: Facets of personal digital document management. *CHINZ 2005*. July 6-8, 2005, Auckland, NZ.

Huvila, I. (2009). Analytical information horizon maps. *Library & Information Science Research, 31*(1), 18-28.

Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation, 52*(1), 3-50.

Isen, A. M. (2004). Positive affect and decision making. In M. Lewis & J. M. Haviland (Eds.), *Handbook of Emotions* (pp.417-35). New York: Guildford.

Jacoby, J., Speller, D., & Berning, C. (1974). Brand choice behavior as a function of information load: Replication and extension. *Journal of Consumer Research, 1*, 33-42.

James, L. and Nahl, D. (1996). Achieving focus, engagement, and acceptance: Three phrases of adapting to Internet use. *EJVC: Electronic Journal of Virtual Culture, (4)*1. Retrieved from http://www2.hawaii.edu/~nahl/articles/ejvc.html on January 27, 2007.

Janes, J. (2003). What is reference for? *Reference Services Review, 31*(3), 22-25.

Japzon, A. (2008). A model of and support for a digital preservation infrastructure that connects individuals to libraries. Proceedings of the World Library and Information Congress: 74th General Conference and Council. http://www.ifla.org/IV/ifla74/papers/084-Japzon-en.pdf.

Japzon, A. & Gong H. (2005). A neighborhood analysis of public library use in New York City. *The Library Quarterly, 75*, 446-463.

Jiao, Q. & Onwuegbuzie, A. (1997). Antecedents of library anxiety. *Library Quarterly*,*67*, 372-389.

Jones, W. (2004). Finders, Keepers? The present and future perfect in support of personal information management. *firstmonday 9*(3). Retrieved on February 21, 2007 from http://www.firstmonday.org/issues/issue9_3/jones/index.html.

Jones, W. (2007). Personal information management. *ARIST, 41,* 453-504.

Jones, W., Phuwanartnurak, A. J., Gill, R., & Bruce, H. (2005). Don't take my folders away!: Organizing personal information to get things done. *Conference on Human Factors in Computing Systems* (pp. 1505-1508). New York, NY: Association for Computing Machinery.

Karger, D. & William, J. (2006, January). Data unification in personal information management. *Communications of the ACM, 49*(1), 77-82.

Kaye, J. J., et al. (2006) To have and to hold: Exploring the personal archive. CHI 2006 (pp. 274-285), April 22-27, 2006, Montréal, Québec, Canada.

Kim, J. (2008). Motivating and impeding factors affecting faculty contribution to institutional repositories. Journal of Digital Information, 8(2). Retrieved on May 5, 2009, from http://journals.tdl.org/jodi/article/viewArticle/193/177.

Krathwohl, D. R. (1997). *Methods of Education and Social Science Research*. 2nd Ed. Boston: Allyn & Bacon.

Krikelas, J. (1983). Information-seeking behaviors: patterns and concepts. *Drexel Library Quarterly, 19*, 5-20.

Kuhlthau, C. (1991) Inside the search process: Information seeking from the user's perspective. *JASIS, 42* (5), 361-371.

Kuny, T. (1998). The Digital Dark Ages? Challenges in the preservation of electronic information. *International Preservation News,* No. 17. Retrieved on October 22, 2006 from http://www.ifla.org/VI/4/news/17-98.htm#2.

Kwasnik, B. H. (1989, June). How a personal document's intended use or purpose affects its classification in an office. In *Proceedings of the ACM-SIGIR 12th Annual International Conference on Research and Development in Information Retrieval.* Cambridge, MA, New York: Association for Computing Machinery.

Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics, 19*(1), 55-66.

LeFurgy, W. (2005, Summer). Building preservation partnerships: The Library of Congress National Digital Information Infrastructure and Preservation Program. *Library Trends, 54* (1), 163-72.

Lincoln, Y & Guba, E. (1985). Naturalistic Inquiry. Thousand Oaks, CA: Sage Publications, Inc.

Lyman, P. & Varian, H. (2003). How much information 2003? Retrieved from http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/ on January 3, 2008.

Lynch, C**.** (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *portal: Libraries and the Academy*, 3(2), 327-336.

Marshall, C. (2006). Maintaining personal information: Issues associated with long-term storage, preservation, and access. Retrieved on August 14, 2006 from

http://www.csdl.tamu.edu/~marshall/PIM%20Chapter-Marshall.pdf.

Marshall, C. (2008a). Rethinking personal digital archiving part 2. *D-Lib Magazine, 14*(3/4).

Marshall, C. (2008b). Rethinking personal digital archiving part 1. *D-Lib Magazine, 14*(3/4).

Marshall, C., Bly, S., & Brun-Cottan, F. (2006). The long term fate of our digital belongings: Towards a service model for personal archives. In Proceedings of IS&T Archiving 2006, Ottawa, Canada, May 23-26, 2006. Springfield, VA: Society for Imaging Science and Technology, pp. 25-30.

Martin, S. P. & Robinson, J. P. (2007). The income digital divide: trends and predictions for levels of Internet use. *Social Problems, 54*(1), 1-22.

Mason, I. (2007). Cultural information standards – Political territory and rich rewards. In F. Cameron & S. Kenderdine (Eds.), *Theorizing Digital Cultural Heritage* (pp.223-244). Cambridge, MA: MIT Press.

Mellon, C. (1986). Library anxiety: A grounded theory and its development. *College & Research Libraries, 47,* 160-165.

Miles, M. B. & Huberman, A. M. (1994). Making good sense: Drawing and verifying conclusions. In *Qualitative data analysis: An expanded sourcebook, 2nd Edition* (pp. 245-287). Thousand Oaks, CA: Sage Publications.

Nahl, D. (1997). Information counseling inventory of affective and cognitive reactions while learning the Internet. *Internet Services Reference Services Quarterly,* 2 (2/3), 11-33.

Nahl, D. & Bilal, eds. (2007). *Information and Emotion*. Medford, NJ: Information Today.

National Diet Library of Japan. (2008). *Ensuring long-term preservation and usability of digital information.* Retrieved on May 5, 2009, from http://www.ndl.go.jp/en/aboutus/preservation.html.

OCLC/RLG Working Group on Preservation Metadata. (2002, June). Preservation Metadata and the OAIS information model. A report. Retrieved from http://www.oclc.org/research/projects/pmwg/pm_framework.pdf on June 18, 2007.

Office of Digital Collections and Research, University of Maryland, College Park. (2007). *Best Practice Guidelines for Digital Collections*. Ed. By Susan Schreibman.

PARADIGM. (2008). *Workbook on digital Private Papers*. Retrieved on May 5, 2009, From http://www.paradigm.ac.uk/workbook/index.html.

Petrelli, D., van den Hoven, E., & Whittaker, S. (2009). Making history: Intentional capture of future memories. *CHI 2009* (pp.1723-1732), April 4-9, 2009, Boston, Massachusetts.

Porck, H. & Teygeler, R. (2000, December). *Preservation Science Survey*. Retrieved on May 5, 2009, from http://www.clir.org/pubs/reports/pub95/contents.html.

Rasmussen, (1980). The human as a systems component. In Smith and Green (eds.), *Human Interactions with Computers*. London: Academic Press.

Reitz, J. (2007). *ODLIS: Online Dictionary of Library and Information Science*. Westort, CT. Libraries Unlimited. Retrieved from http://lu.com/odlis/about.cfm on January 3, 2008.

Rodden, K. & Wood K. (2003). How do people manage their digital photographs? *CHI 2003*. April 5-10, 2003. Ft. Lauderdale, Florida, USA.

Ross, S. & Hedstrom, M. (2005). Preservation research and sustainable digital libraries. *International Journal on Digital Libraries, (5)*, 317-324.

Rothenberg, J. (1998). Avoiding technological quicksand: Finding a viable technical foundation for digital preservation. A Report from the Council on Library and Information Resources. Retrieved from http://www.clir.org/PUBS/reports/rothenberg/contents.html on July 31, 2006.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R.J. Spiro, B.Bruce, & W.F. Brewer (eds.), *Theoretical Issues in Reading and Comprehension*. Hillsdale, NJ: Erlbaum.

Savolainen, R. (1995). Everyday life information seeking: approaching information seeking in the context of 'Way of Life.' *LISR, 17* (3), 259-294.

Savolainen, R. & Kari, J. (2004). Placing the internet in information source horizons. A study of Information seeking by internet users in the context of self-development. *LISR, 26*, 415-433.

Savolainen, R. (2006a). Spatial factors as contextual qualifiers of information seeking. Information Research 11(4). Retrieved on March 28, 2007 from http://informationr.net/ir/11-4/paper261.html.

Savolainen, R. (2006b). Time as context of information seeking. *LISR, 28*, 110-127.

Savolainen, R. (2007). Information source horizons and source preferences of environmental activists: A social phenomenological approach. *JASIS&T, 58*(12), 1709-1719.

Savolainen, R. (2008). Source preferences in the context of seeking problem-specific information. *Information Processing and Manage, 44*(1), 274-293.

Schamber, L. (1996). What is a document? Rethinking the concept in uneasy times. *JASIS*, *47*(9), 669-671.

Searle, S. & Thompson, D. (2003, April). Preservation metadata. *D-Lib Magazine, (9)*4. Retrieved from http://www.dlib.org/dlib/april03/thompson/04thompson.html on April 2, 2007.

Sonnenwald, D. H. (1999). Evolving the perspectives of human information behavior: Contexts, situations, social networks and information horizons. In T. Wilson & D. Allen (Eds.), *Exploring the Contexts of Information Behavior* (pp. 176-190).

Sonnenwald, D. H. & Wildemuth, B. M. (2001). A research method to investigate information seeking using the concept of information horizons: an example forma study of lower socio-economic students' information seeking behavior. *The New Review of Information Behavior Research, 2,* 65-86.

Spink, A., Greisdorf, H., & Bateman, J. (1998). From highly relevant to not relevant: Examining different regions of relevance. *Information Processing & Management, 34*(5), 599-621.

Tang, R. & Solomon, P. (1998). Toward an understanding of the dynamics of relevance judgment: An analysis of one person's search behavior. *Information Processing & Management, 34*(2-3), 237-256.

Teevan, J., et al. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. *Conference on Human Factors in Computing Systems* (pp. 415-22). New York, NY: ACM.

Vavrek, B. (2000). Is the American public library part of everyone's life? *American Libraries, 31*, 60-63.

Russo, A. & Watkins, J. (2007). Digital cultural communication: Audience and Remediation. F. Cameron & S. Kenderdine (Eds.), *Theorizing Digital Cultural Heritage* (pp. 149-164). Cambridge, MA: MIT Press.

Whittaker, S. & Sidner, C. (1996). Email overload: exploring personal information Management of email. *CHI 96*. April 13-18, 1996, Vancouver, BC, Canada.

Whittaker, S. & Hirschberg, J. (2001). The character, value, and management of personal paper archives. *ACM Transactions on Computer-Human Interaction, 8*(2), 150-170.

Whittaker, S., Bellottti, V., & Gwizdka, J. (2006). Email in personal information management. *Communications of the ACM, 49*(1), 68-73.

Williams, P., et al. (2008, April). Digital Lives: Report of Interviews with the Creators of Personal Digital Collections. *Ariadne, 55*. Retrieved on May 6, 2009, from http://www.ariadne.ac.uk/issue55/williams-et-al/.

Wilson, T. D. (1997). Information behavior: An interdisciplinary perspective. *Information Processing and Management, 33*(4), 551-572.

Wilson, T. D. (2000). Human information behavior. Informing Science, 3(2), 49-55.

Workshop on Research Challenges in Digital Archiving and Long-term Preservation. (2004). It's about time: Research challenges in digital archiving and long-term preservation, II. *Microform & Imaging Review, 33*(1), 23-36.

Wurman, R. (1989). Information Anxiety. New York: Doubleday.

Wurman, R., Sume, D. & Loring, L. (2000). Information Anxiety 2. Que.

Yakel, E. (2004, October). Seeking Information, seeking connections, seeking meaning: genealogists and family historians. *Information Research, 10*(1). Retrieved from http://informationr.net/ir/10-1/paper205.html on February 11, 2007.

Xu, Y. & Chen Z. (2006). Relevance judgment: What do information users consider beyond topicality? *JASIS&T, 57*(7), 961-973.

Yee, S. (2007). The archive. In *Evocative objects: Things we think with*, ed. S. Turkle. Cambridge, MA: The MIT Press. Pp. 30-37.

Zach, L. (2006). Using multiple-case studies design to investigate the information seeking behavior of arts administrators. *Research Methods, 55*(1), 4-21.

Appendix A: Example of Matrices (Oscar)

Digital Content Stored in Digital Devices

	Desktop/s	Laptop/s	PDA/s	Mobile Phone	Digital Camera/s	iPod/MP3 player	Flash Drive	External Hard Drive	CDRom	DVD	Floppy Disk 3"	Other	Other	Other
(Specify Device Brand or Other)		Mac	Palm		Canon	Apple		Apple						
Database files (e.g. Access)		X						X	X	X				
Digital Video		X	X	X				X	X	X				
Email		X						X	X	X	X			
Music (MP3s or itunes)		X				X		X	X	X				
Other Recordings		X						X	X	X				
Original HTML documents														
PDFs of articles		X						X	X	X				
Personal Word Processed Documents		X						X	X	X	X			
Photographs		X	X	X				X	X	X				
Images other than photographs		X						X	X	X				
Presentation (PowerPoint) Slides		X						X	X	X				
Scanned personal documents		X						X	X	X				
Scanned personal objects														
Scanned personal photos														
Spreadsheets		X						X	X	X				
Web pages (html docs not original)		X						X	X	X				
Other -Movies								X	X	X				
Other - TV shows		X						X	X	X				
Other - Address Book		X	X	X		X		X	X	X				
Other - Podcasts		X				X		X	X	X				

Personal Digital Content Stored on the Web

	Blog	Face-book	Flickr	Google	Web Browser	MySpace	Personal Website	School/ company server space	Snap-fish	Shutter-fly	YAHOO	YouTube	Collabor-ative Spac	Other
(Specifiy Vendor, Feature or Other)														.mac
Bookmarks					X			X						X
Database files														X
Digital Videos														
Email				X				X						X
Music (MP3s or itunes)														X
Other Recordings														
Original HTML documents														
PDFs of articles														
Personal Word Processed Documents								X						X
Photographs														X
Images other than photos														
Presentation (PowerPoint) Slides														X
Scanned personal documents														
Scanned personal objects														
Scanned personal photos														
Spreadsheets														X
Web pages (html docs not original)														
Other - Address Book														X
Other - Calendar														X
Other -														

Appendix B: Interview Guide

Introduction

Thank you for agreeing to participate in this research regarding how people value personal information such as letters, emails, photographs, journals, etc.
I am seeking to understand how individuals use and store personal information and how that might be changing in the digital environment. For the next hour or so I will ask you questions about you and your own personal information. Your answers will provide valuable insights for my continuing research project.

I. Information General

A. Discuss the different types of information that you keep or collect that are of personal importance or interest to you. How long you've been doing so for each type?

B. For information that you have in both physical and digital forms, do you value you one more than the other. Or just differently?

C. Describe one piece of personal information that is physical in form and one piece of information that is digital that you value greatly? Why is each valuable to you? How do you care for those possessions? How long do you plan to keep each one?

II. PIM

A What types of personal information do you save on your computer, other personal digital devices, and or web spaces like Flickr or Snapfish? Why and how do you save them? (Have participant fill out personal information matrices as part of this question.)

B. What types of information do you use most frequently? Why?
How do you access items you use most frequently? Where are they stored?

C. Do you treat personal digital information that you would like to preserve over many years differently than personal digital information you plan to only use for the short term? If so how so?

D. Do you back up all the content on your computer and/or digital devices?

E. What other steps do you take to preserve your digital possessions over the long term?

F. How do you find content that you have saved on your computer or other digital devices?

III. Incident –Successful

A. Please describe a positive experience with a digital information item, an experience that was joyful, creative or memorable?

B. Start by discussing how you came to have the digital item in the first place?

Why you decided to save or create it?

C. What format did you save it in and why? Do you believe this format to be a good one for preserving this item over the long term?

D. Are you pleased with the format? Why or why not?

IV. Incident – Unsuccessful

A. Please describe a negative experience with digital information, perhaps you lost a digital photo or document?

B. How did this loss or experience make you feel?

C. Since this happened have you changed the way you save these kinds of items? If so, how?

D. Thinking generally about using your computer and other digital devices to view or hear digital possessions, do you find that having to experience digital content via a computer screen and/or by some software application enhances or detracts from your enjoyment of digital content? Ask for experiences.

V. Information Seeking

A. Think of a time recently when you needed information regarding your computer, a digital device a software program or your digital content. What type of information did you need and why did you need it? Who did you go to for help or what resources did you use? Were you satisfied with the information you received. Would you do the same the next time you needed this type of information or would you do something differently?

B. What is the best way for you to learn about computer or digital technology?

VI. Information Source Horizon Map.

The participants will be asked to fill in the three zones for the following types of personal information sources.

1. Personal information the participant values across all types of situations and for all time periods. The participant will be asked to explain the level of value for each zone.
2. Personal information the participant wishes to preserve over the long term, at least 5 years, but as long as 25 or indefinitely. The participant will be asked to explain the level of value for each zone.

VII. Conclusion Is there something you would like to add? Is there something you wished I had asked you about but did not?

Appendix C: Instructions for Filling in the three zones of the Maps

First Map – Information valuable to everyday life

Please think about the information, print and/or digital, that you consider across all situations in everyday life, focusing on information of value to your personal life rather than your professional work life. In the center zone, please indicate the information of most **valuable to you in your daily life.** In the middle zone, indicate the next most valuable, and in the outermost zone the next most. In other words, the information should decrease in value as you move towards the outer zone. Once you have completed filling in the zones, I am going to ask you a few questions about the criteria you used to determine the value of each zone.

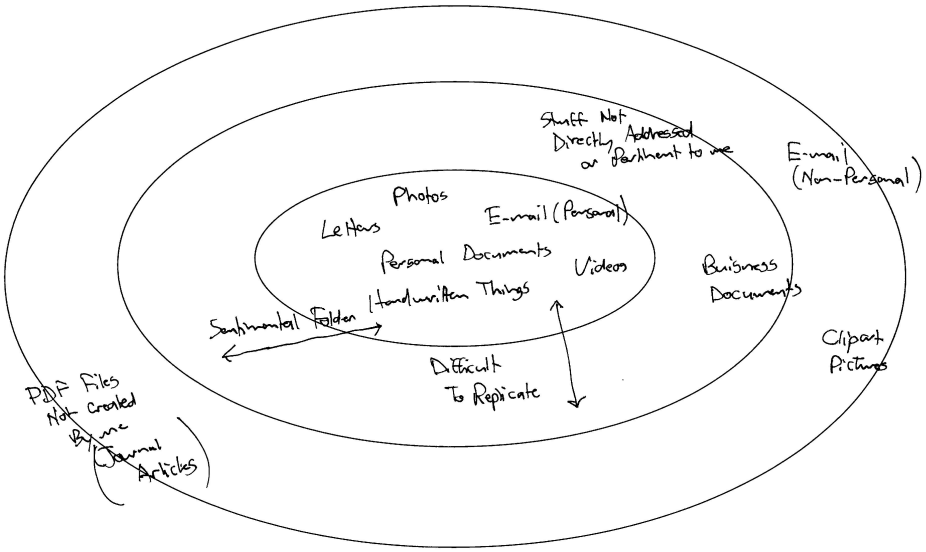
Second Map – Information valuable over the long term

Please think about the information, print and/or digital, that you **wish to preserve over the long term, at least 5 years, as long as 25, or indefinitely.** In the center zone, please indicate the information of most value to you for long term preservation. In the middle zone, indicate the next most valuable, and in the outermost zone the next most. In other words, the information should decrease in value as you move towards the outer zone. Once you have completed filling in the zones, I am going to ask you a few questions about the criteria you used to determine the value of each zone.

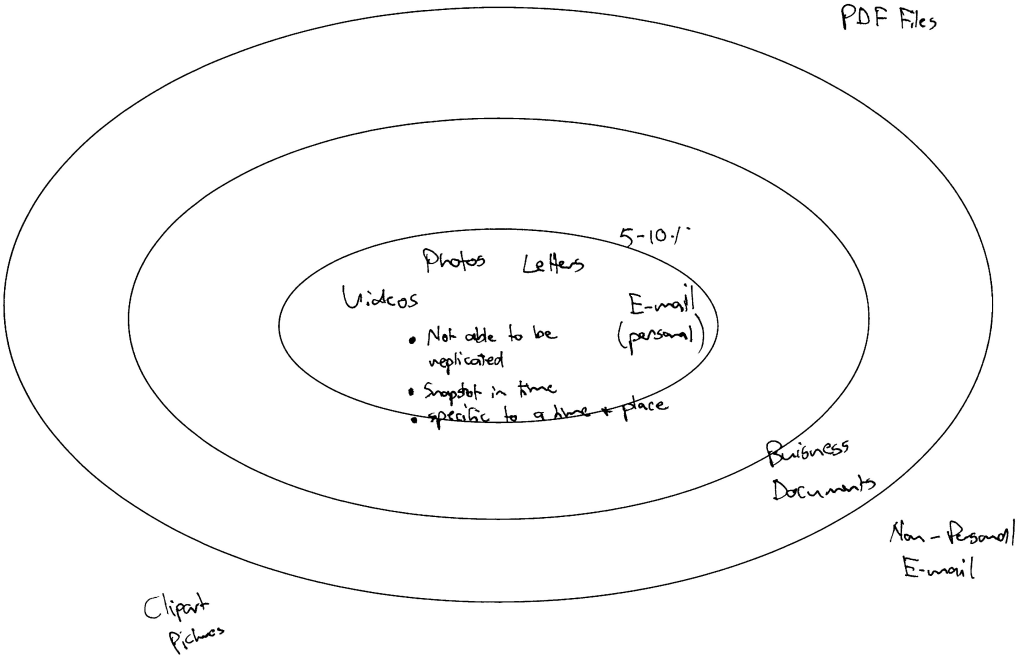
Appendix D: Information Source Horizon Maps For All Participants

Information Source Horizon Maps (Bill)

Everyday

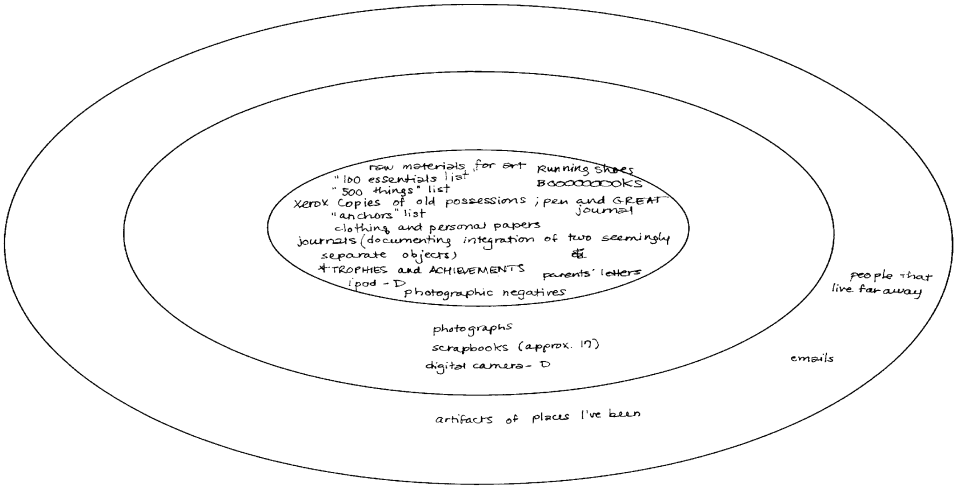


Long Term

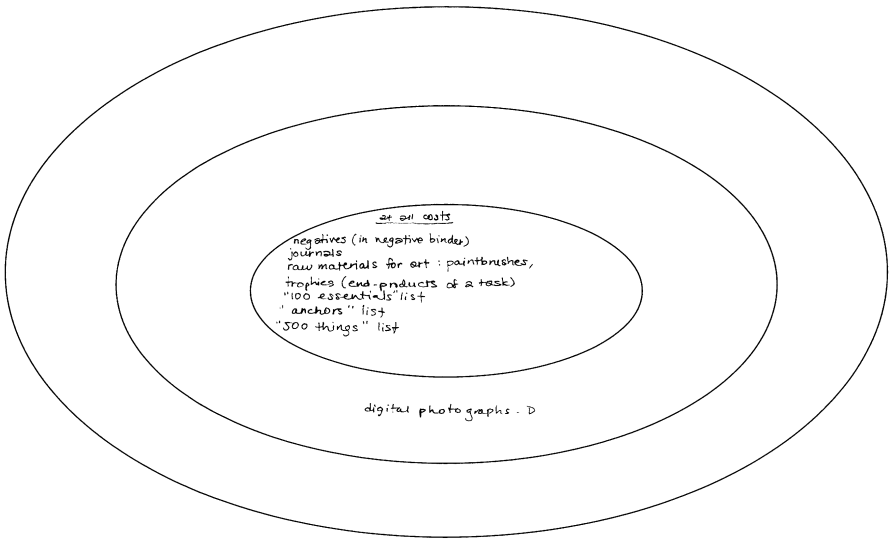


Information Source Horizon Maps (Cathy)

Everyday

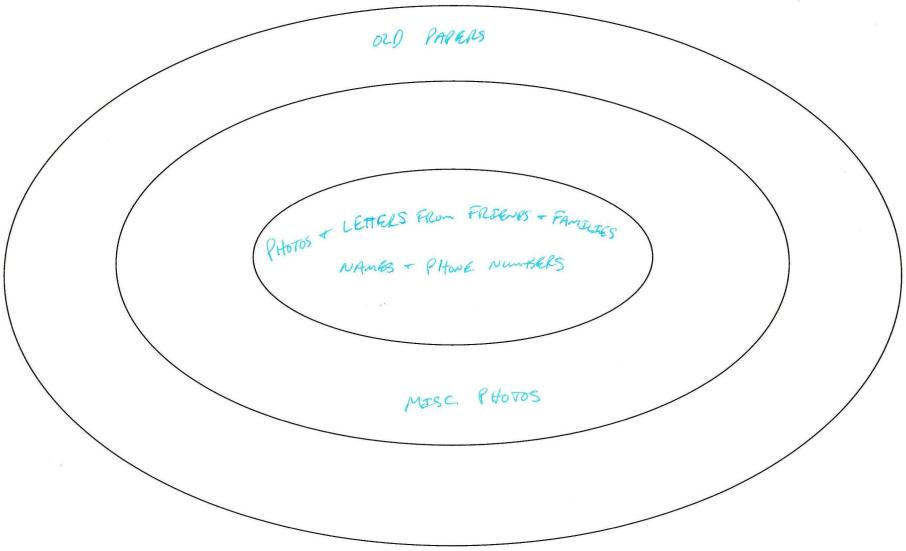


Long Term

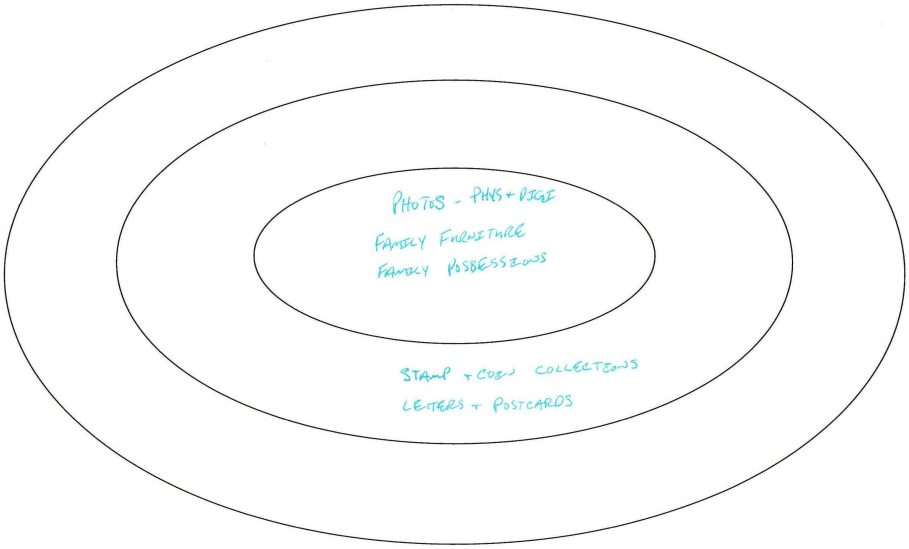


Information Source Horizon Maps (David)

Everyday

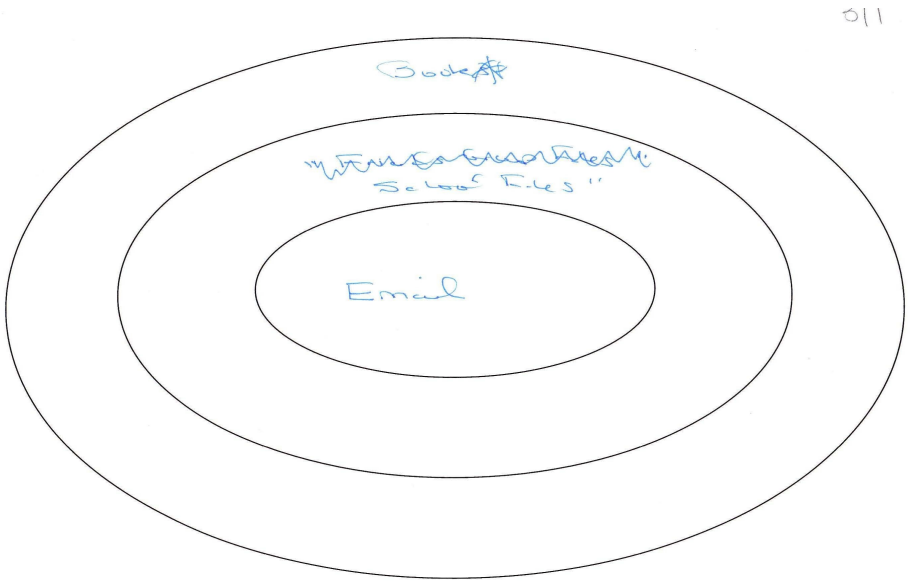


Long Term

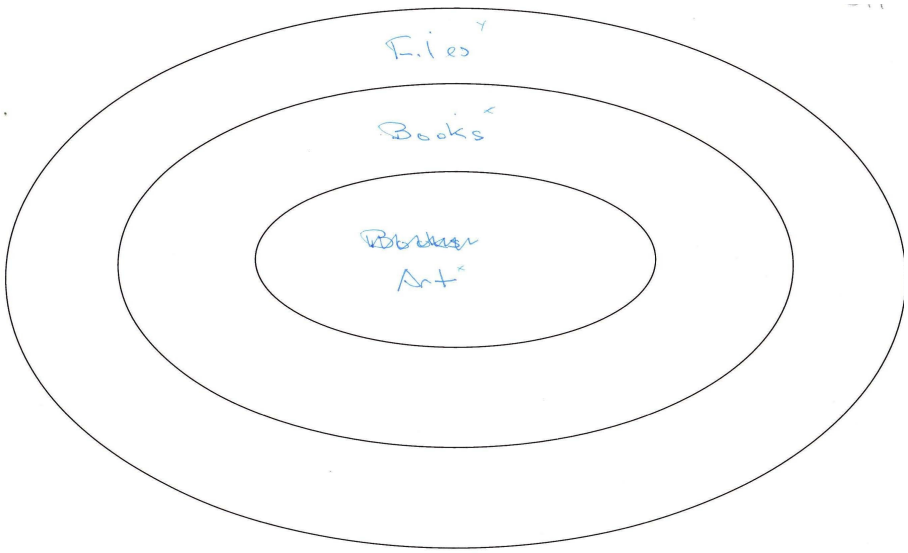


Information Source Horizon Maps (Eleanor)

Everyday

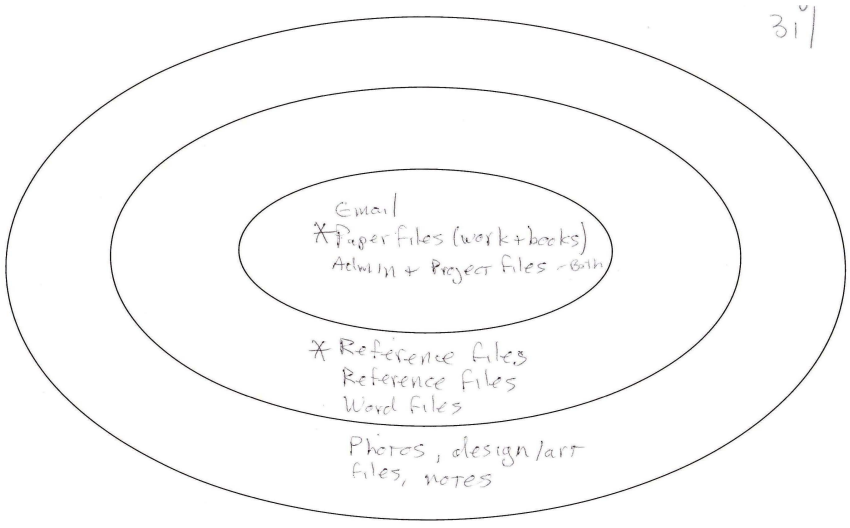


Long Term

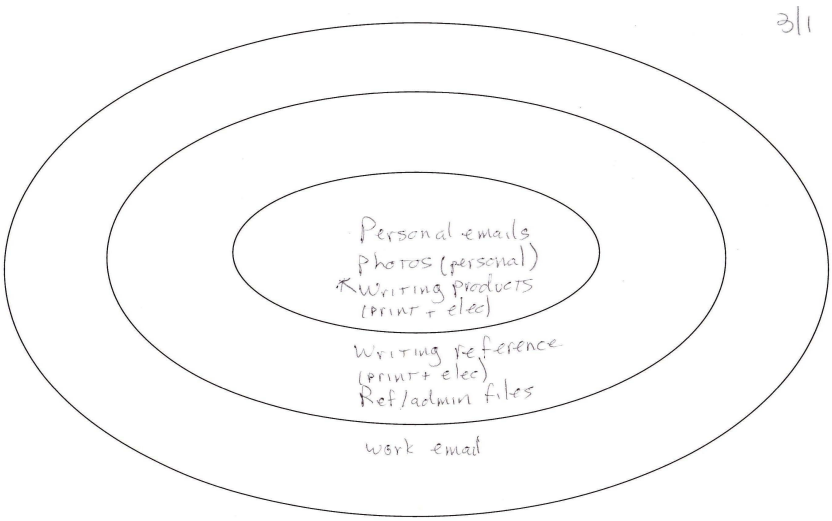


Information Source Horizon Maps (Frank)

Everyday

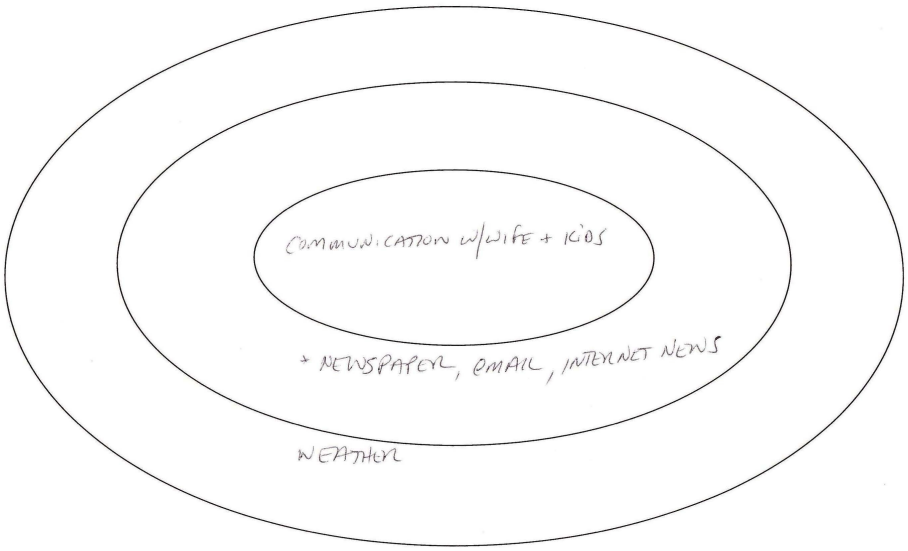


Long Term

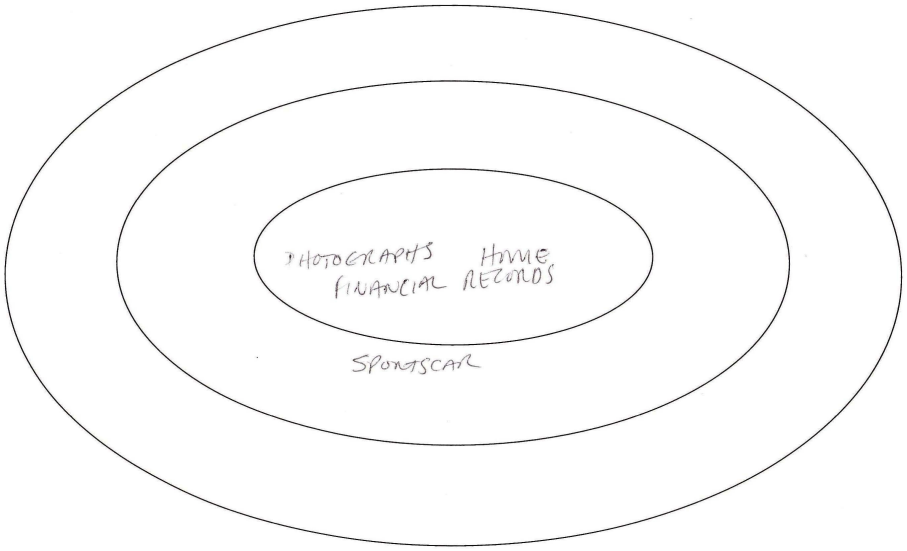


Information Source Horizon Maps (Greg)

Everyday



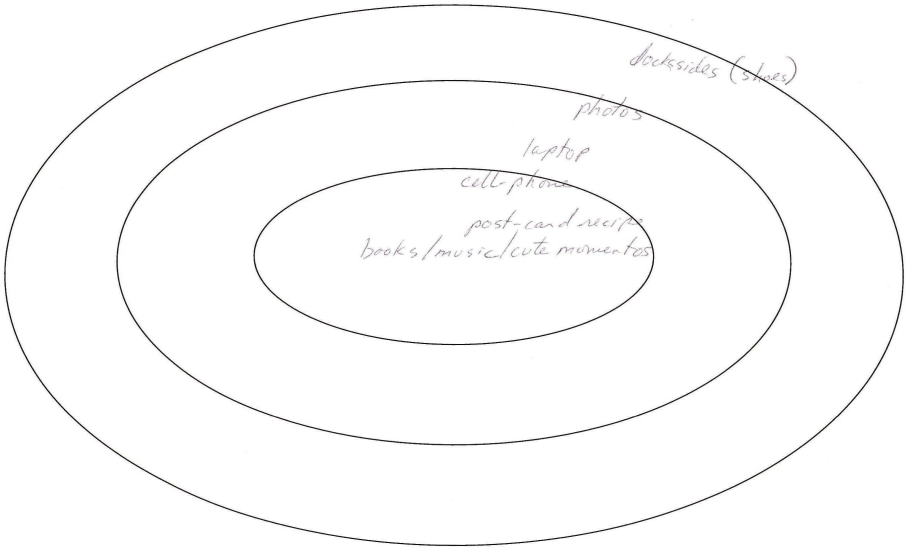
Long Term



Information Source Horizon Maps (Holly)

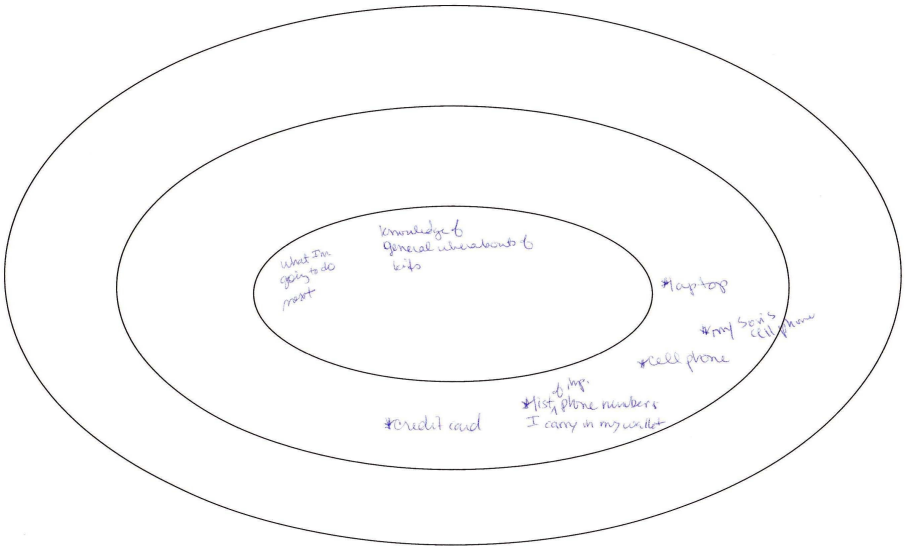
Everyday

Long Term

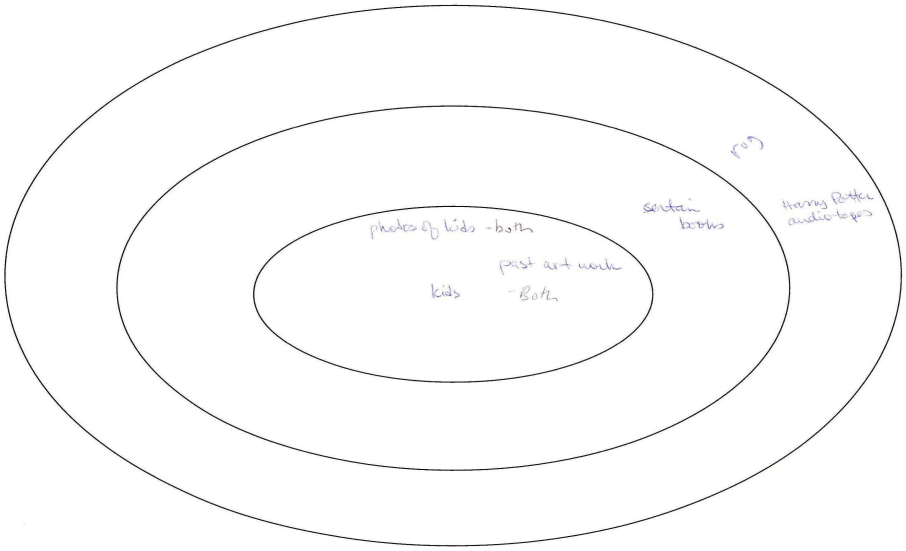


Information Source Horizon Maps (Irene)

Everyday

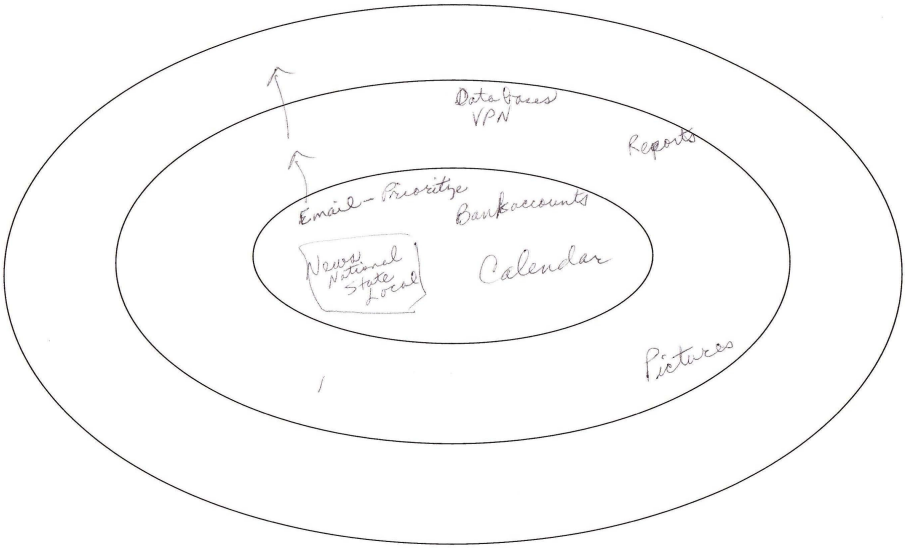


Long Term



Information Source Horizon Maps (Julie)

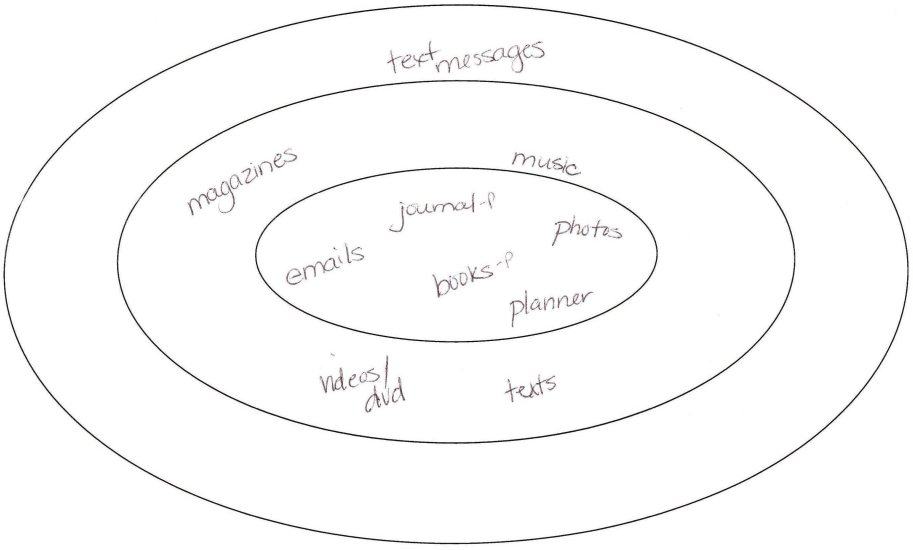
Everyday



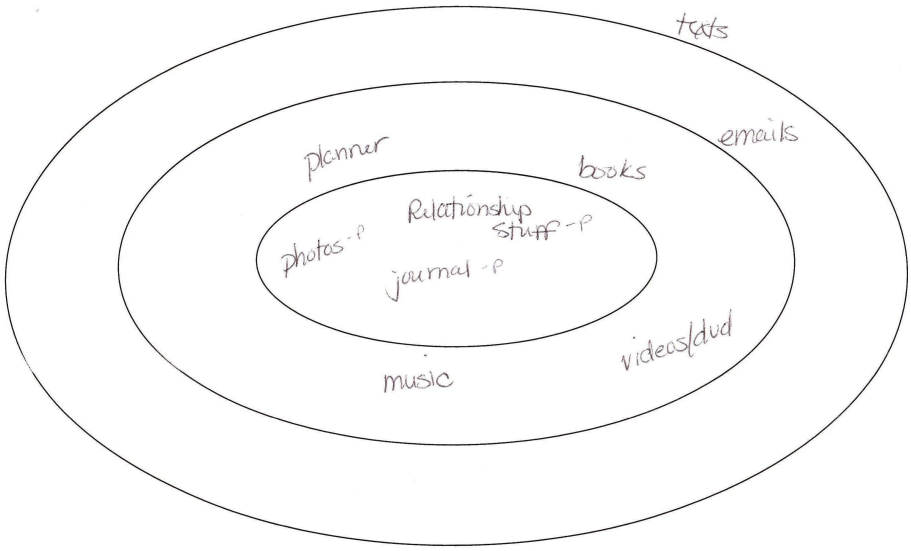
Long Term

Information Source Horizon Maps (Kelly)

Everyday

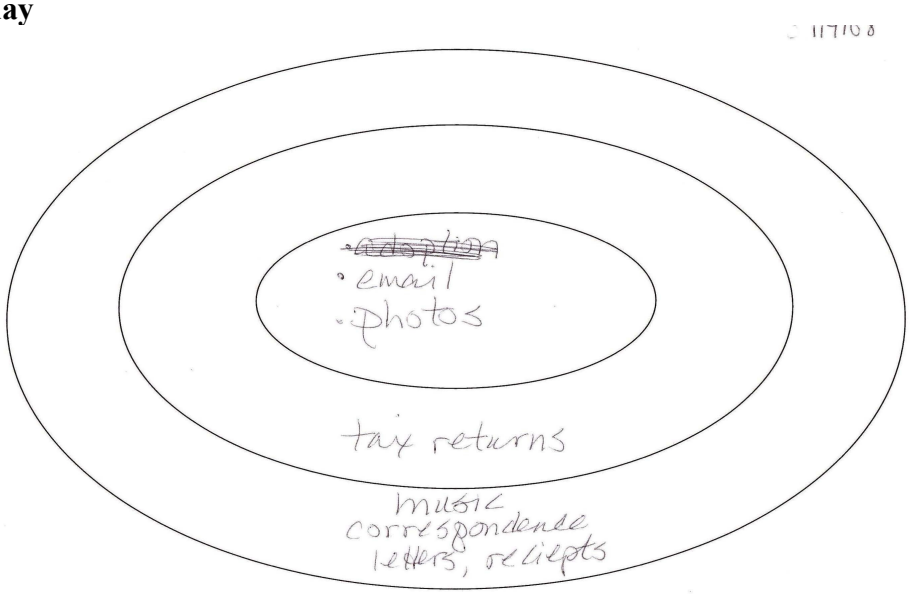


Long Term

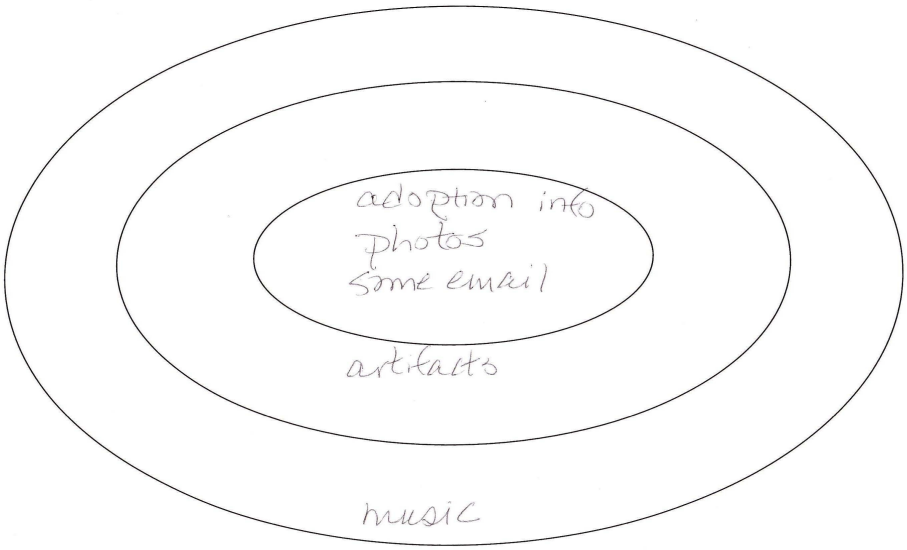


Everyday Information Source Horizon Maps (Lisa)

Everyday

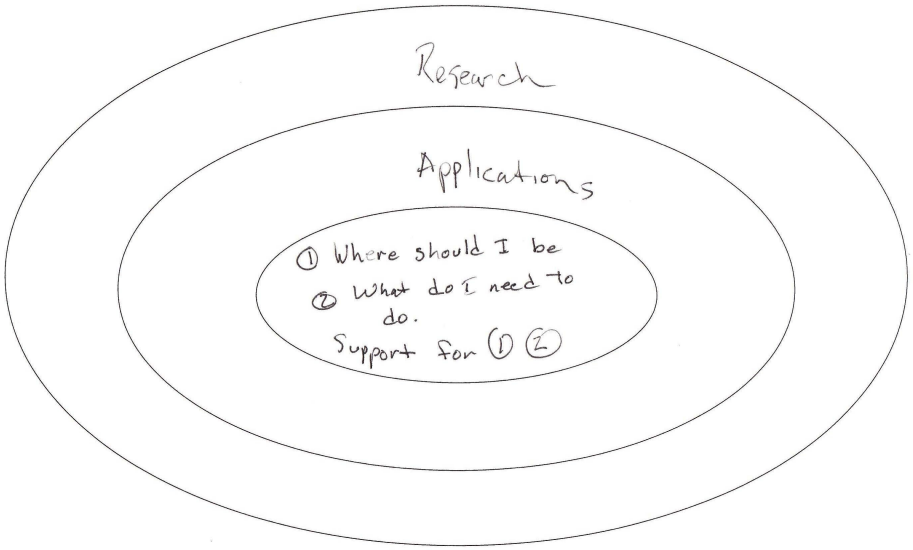


Long Term

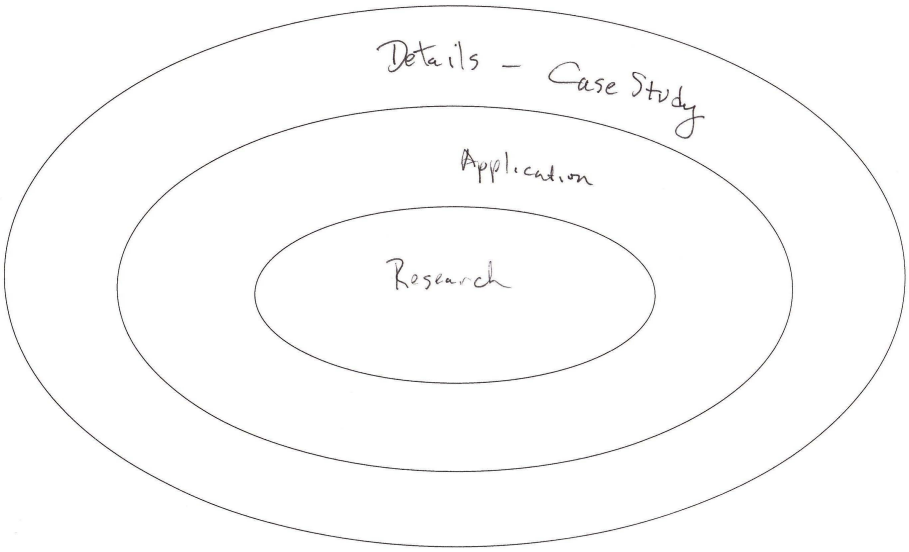


Information Source Horizon Maps (Mark)

Everyday

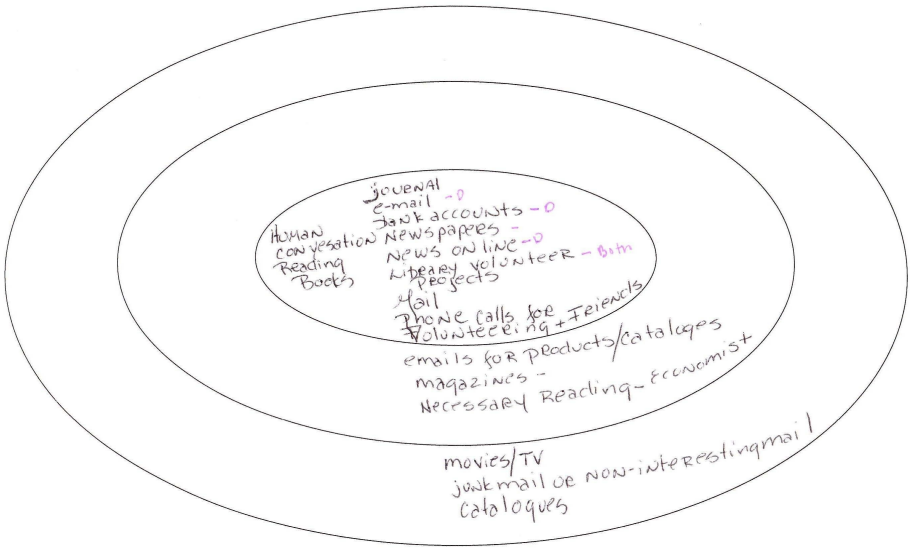


Long Term

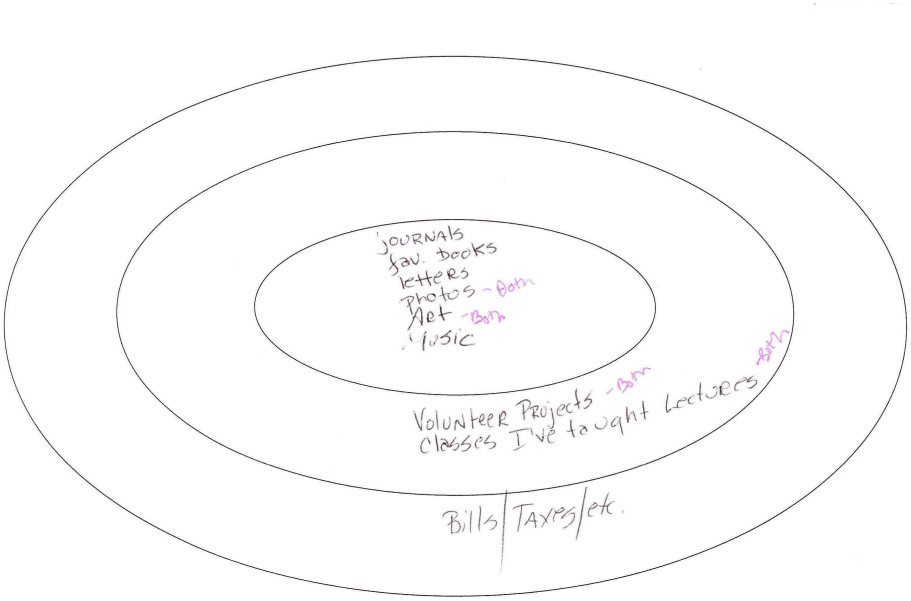


Information Source Horizon Maps (Nora)

Everyday

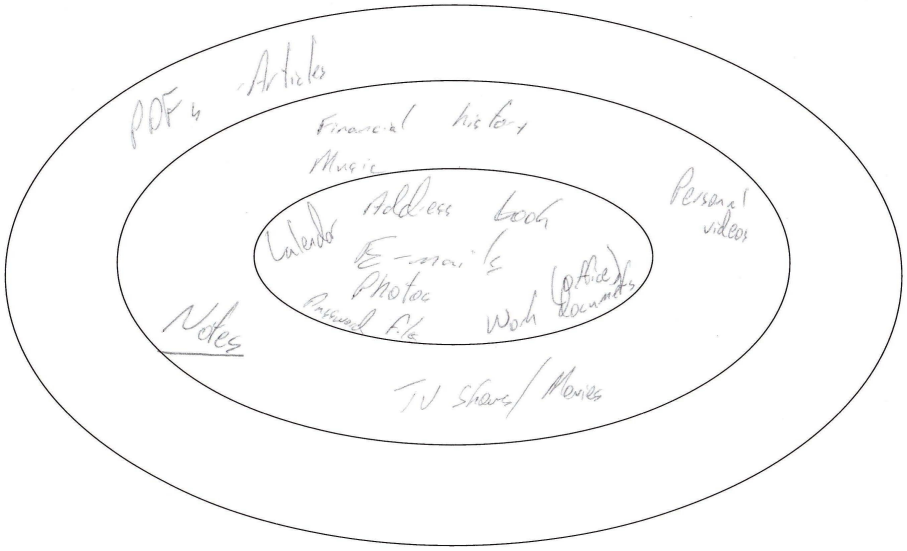


Long Term

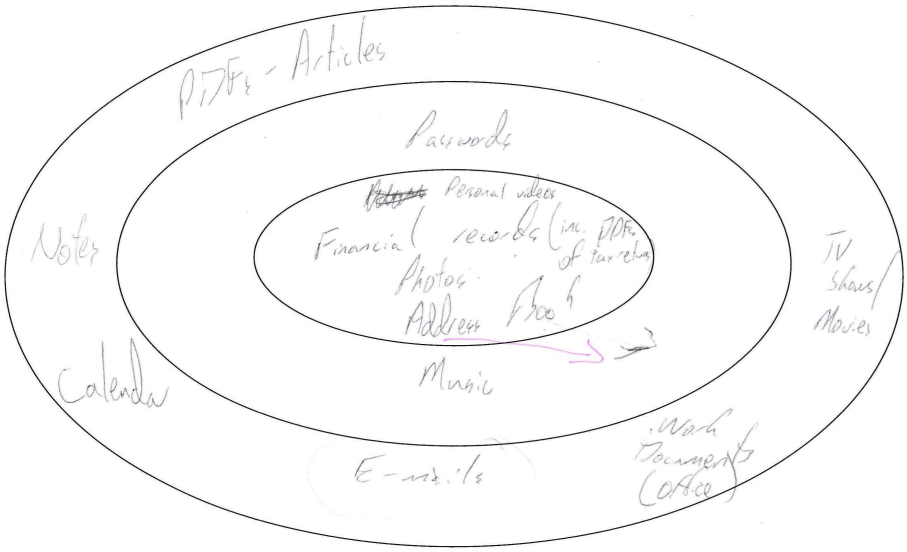


Information Source Horizon Maps (Oscar)

Everyday

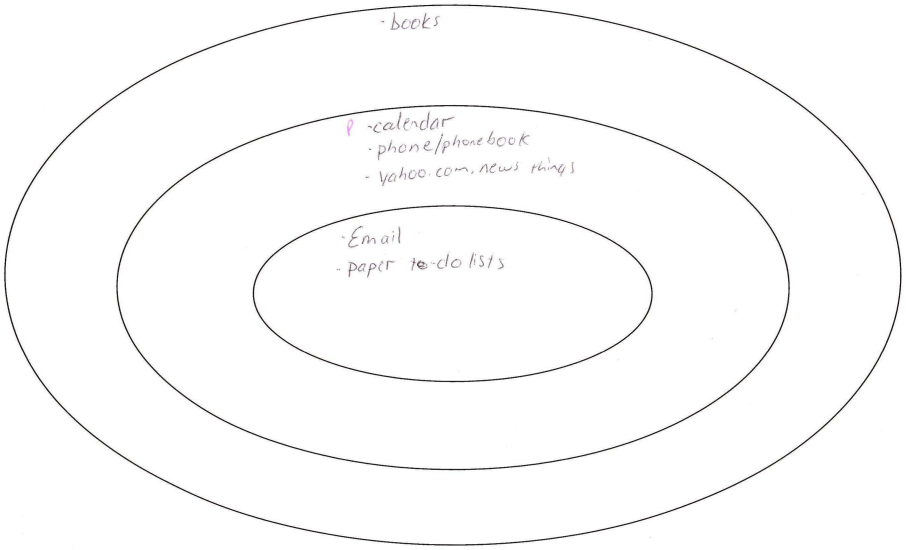


Long Term

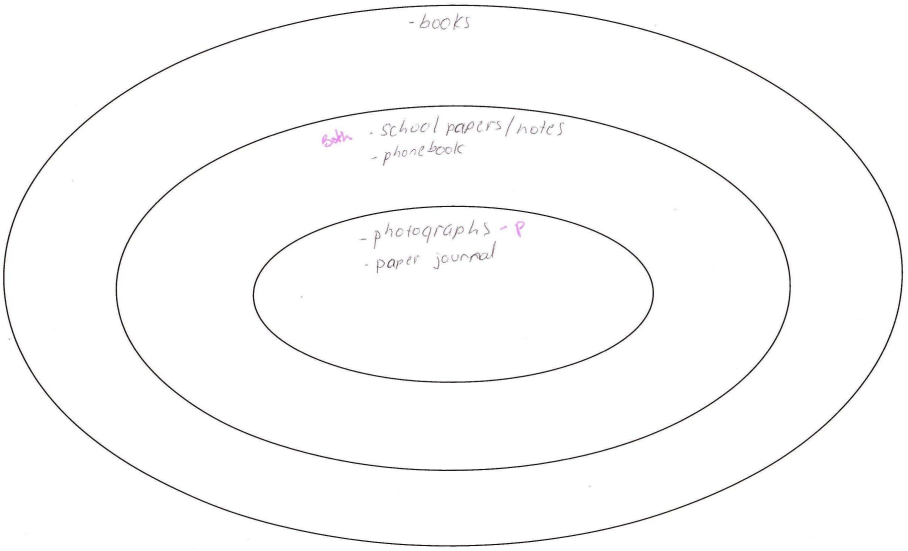


Information Source Horizon Maps (Paula)

Everyday

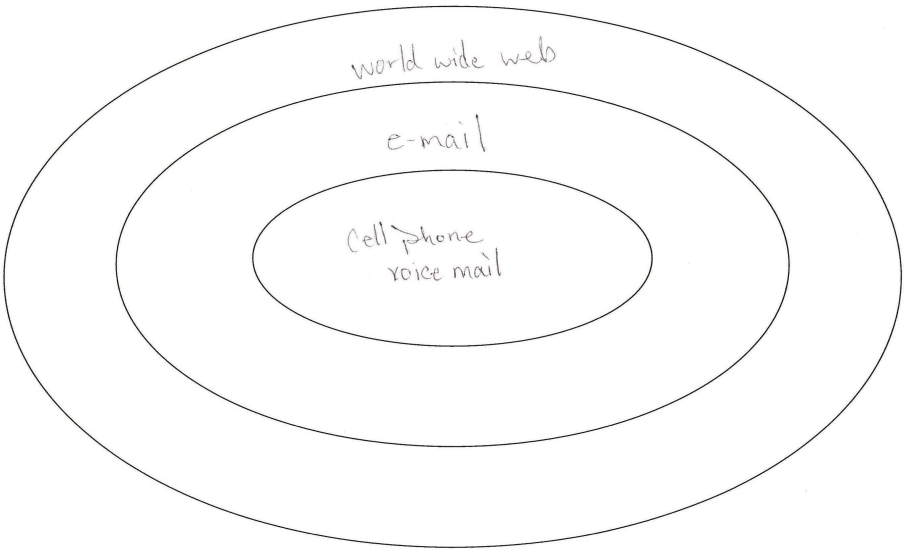


Long Term

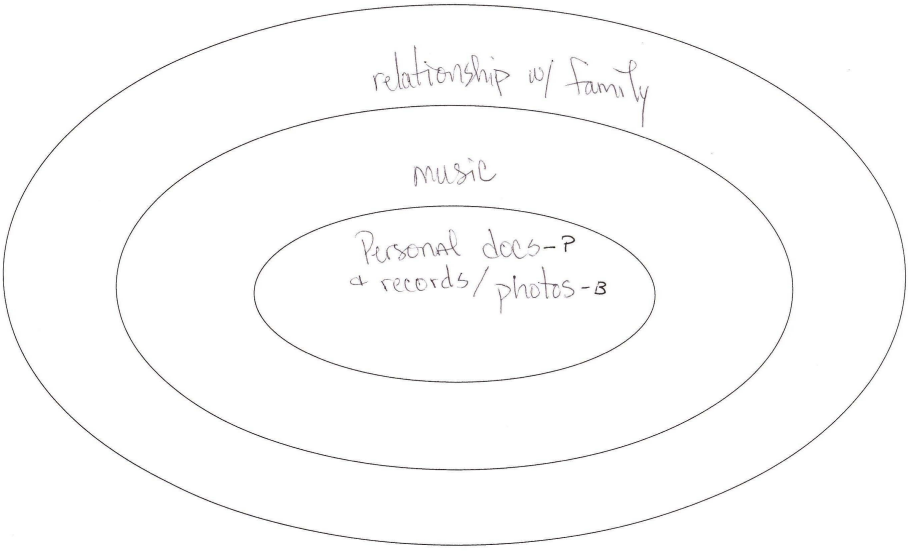


Information Source Horizon Maps (Quincy)

Everyday

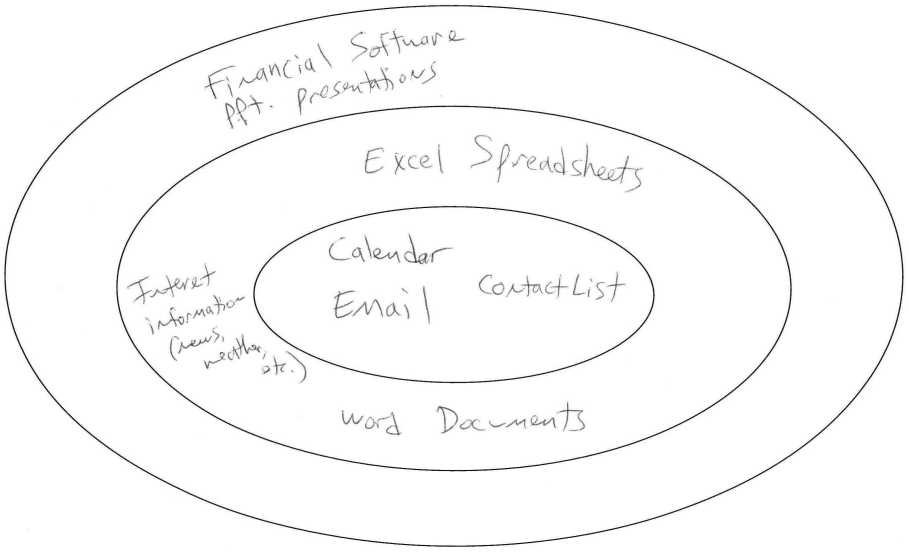


Long Term

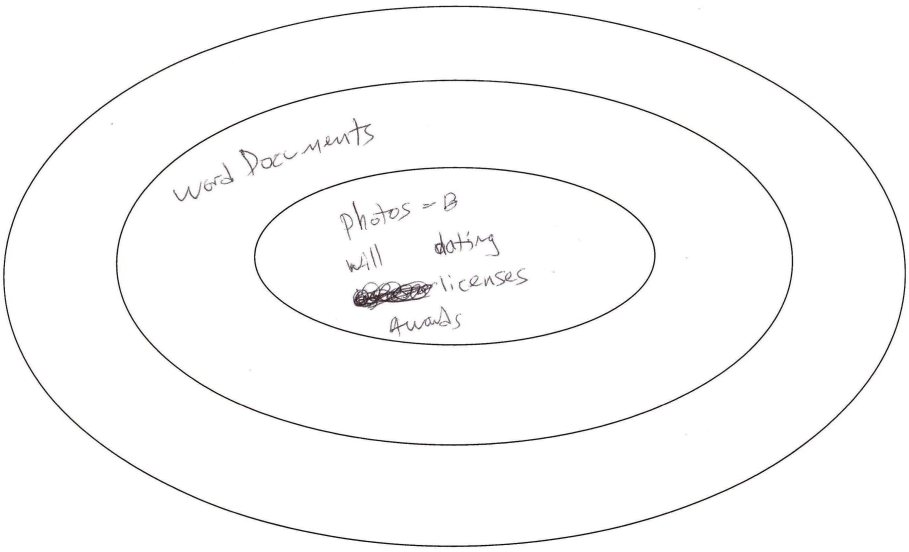


Information Source Horizon Maps (Robert)

Everyday

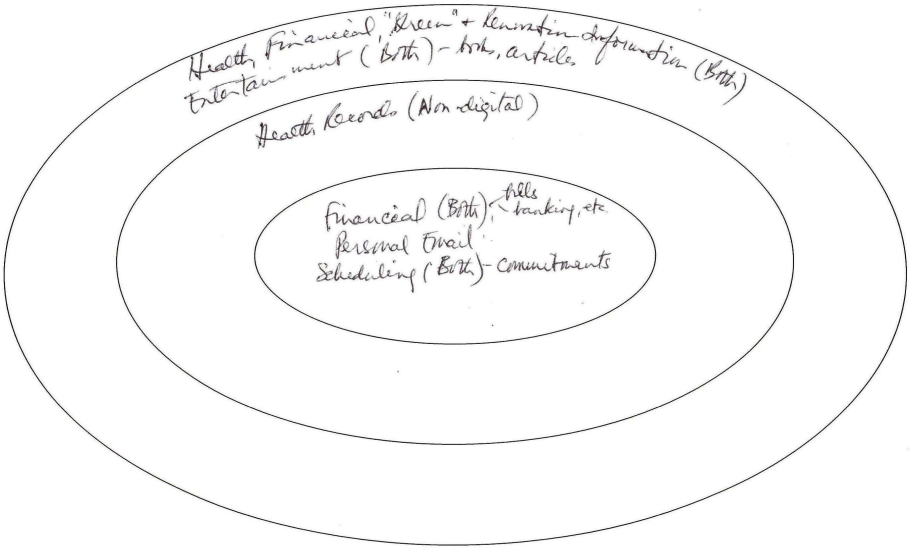


Long Term

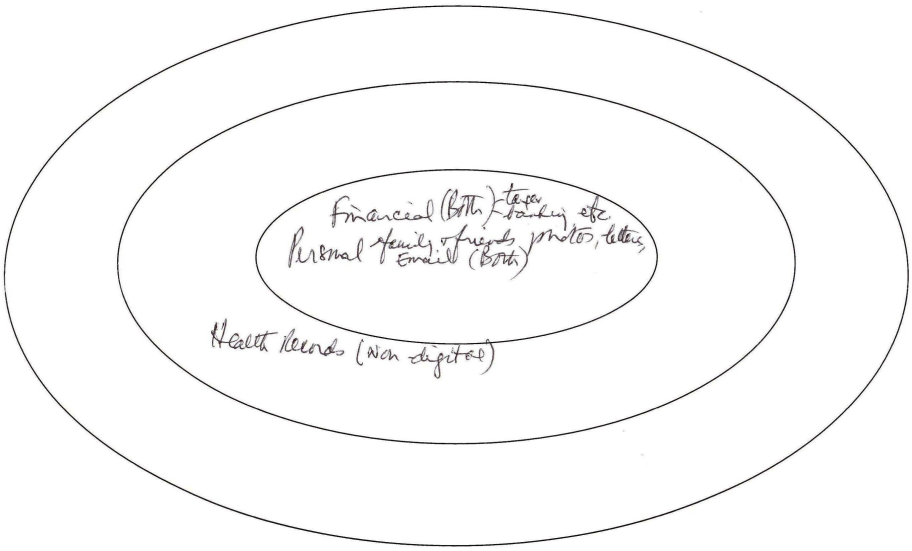


Information Source Horizon Maps (Sarah)

Everyday

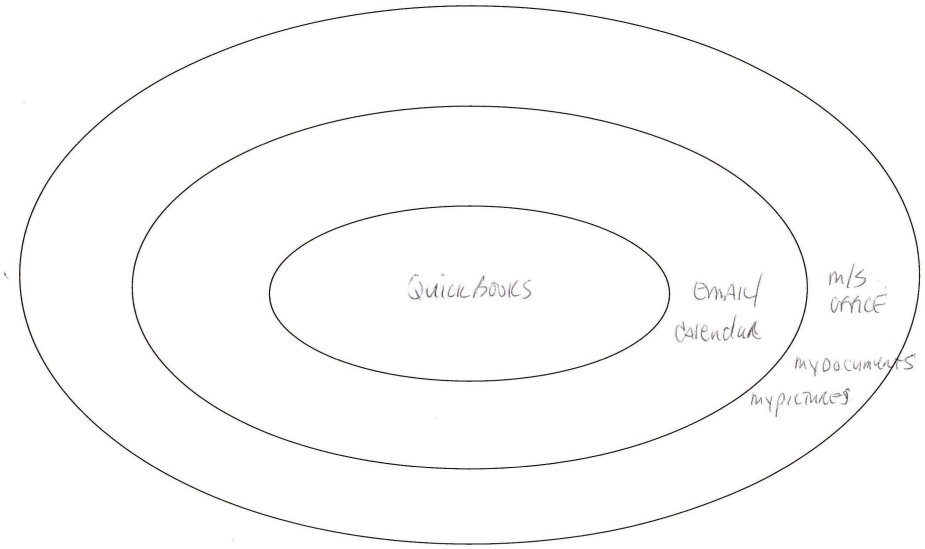


Long Term

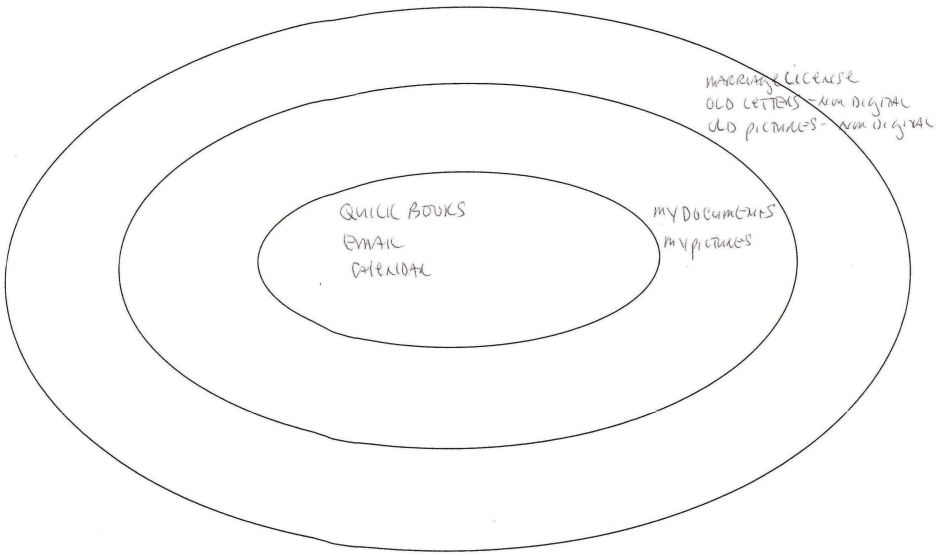


Information Source Horizon Maps (Tina)

Everyday

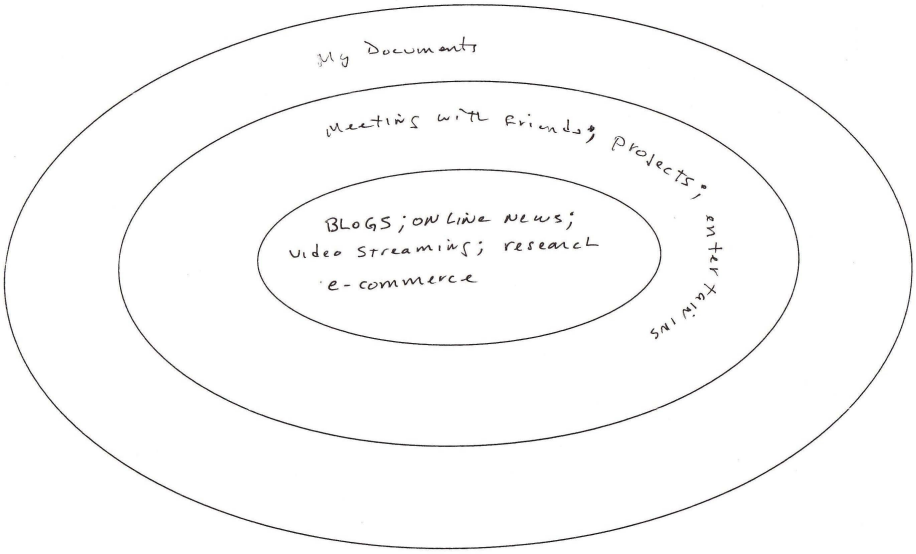


Long Term

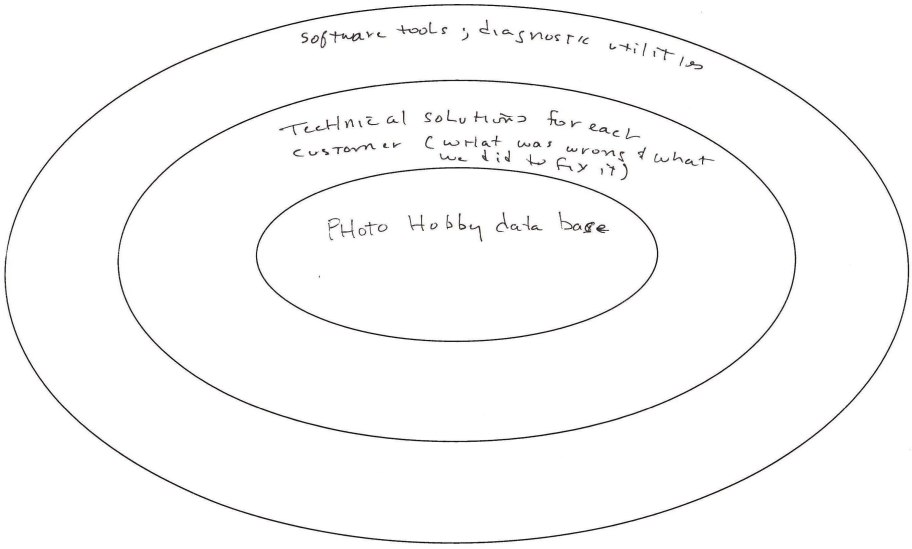


Information Source Horizon Maps (Uri)

Everyday

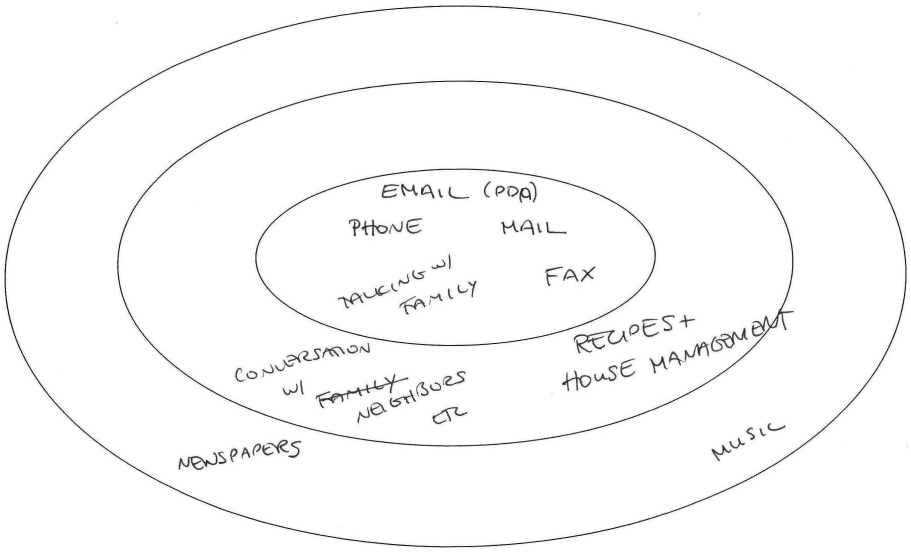


Long Term

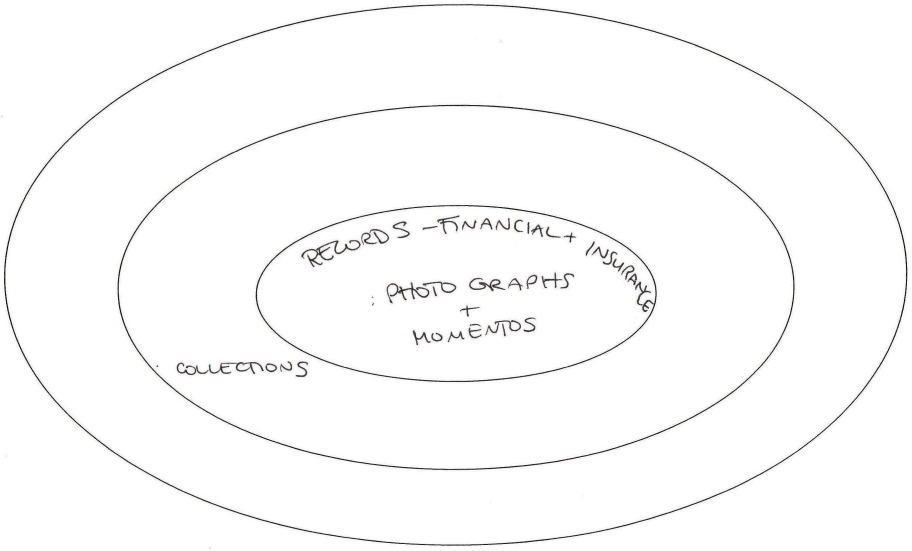


Information Source Horizon Maps (Victoria)

Everyday

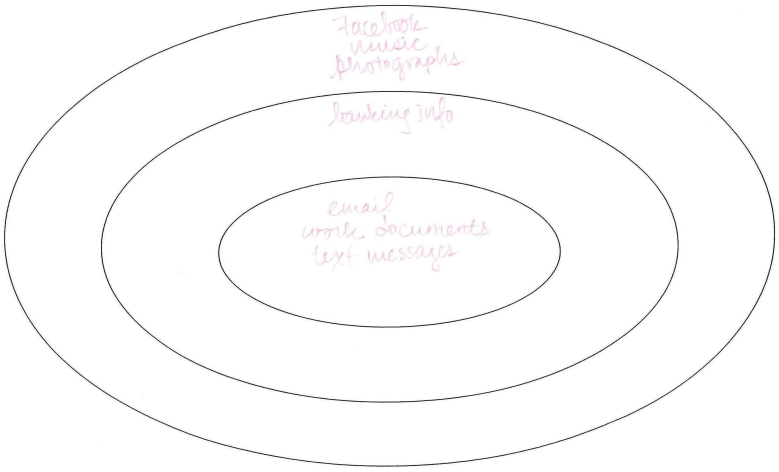


Long Term

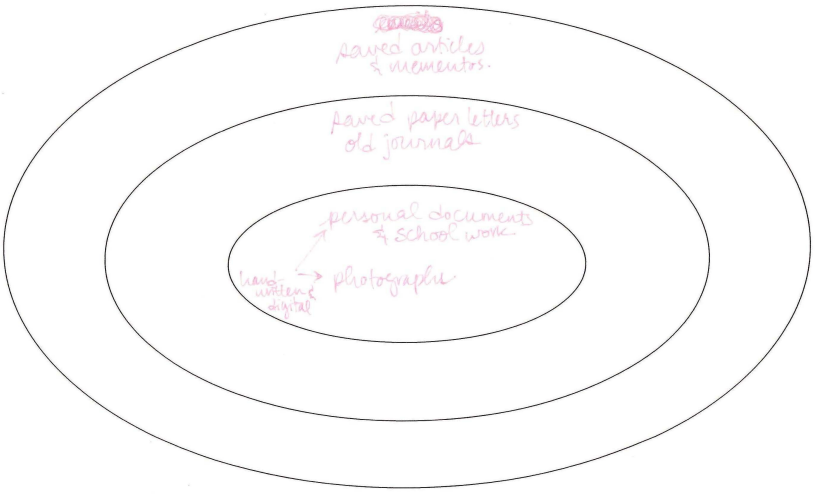


Information Source Horizon Maps (Wendy)

Everyday

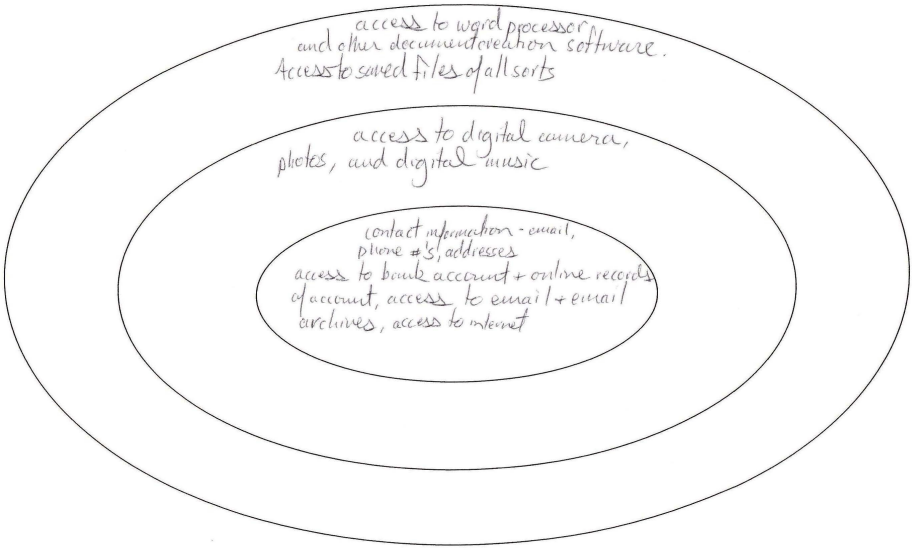


Long Term

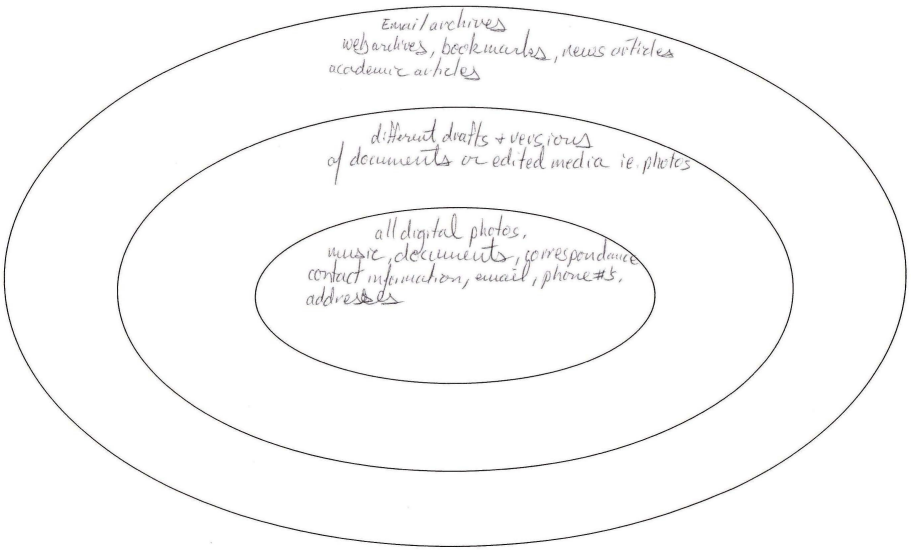


Information Source Horizon Maps (Xavier)

Everyday

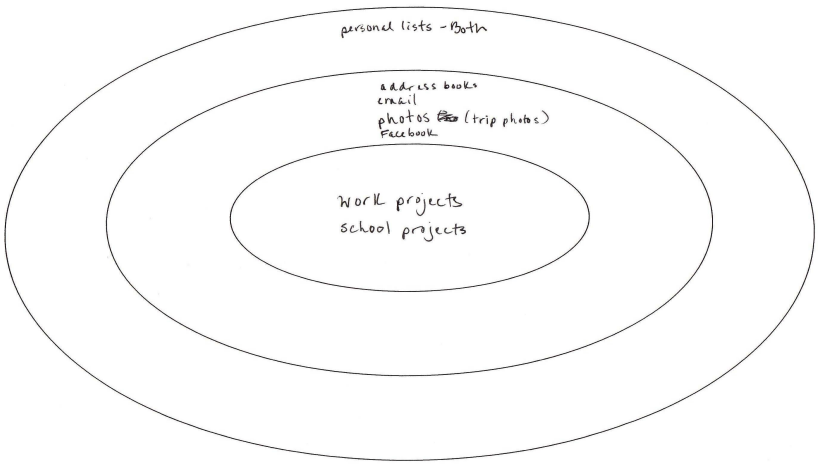


Long Term

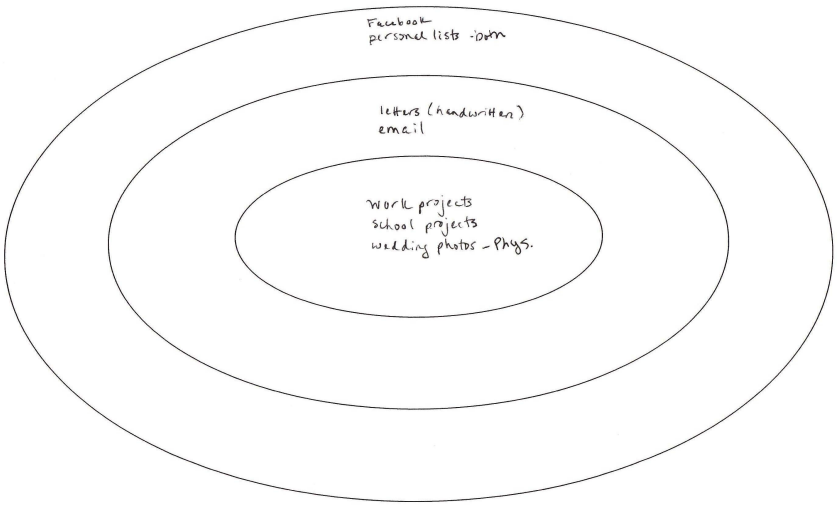


Information Source Horizon Maps (Yvonne)

Everyday

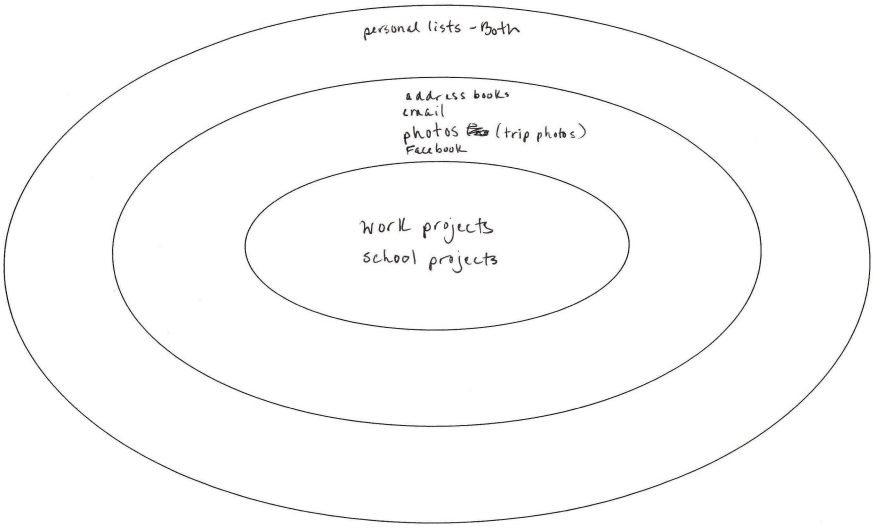


Long Term

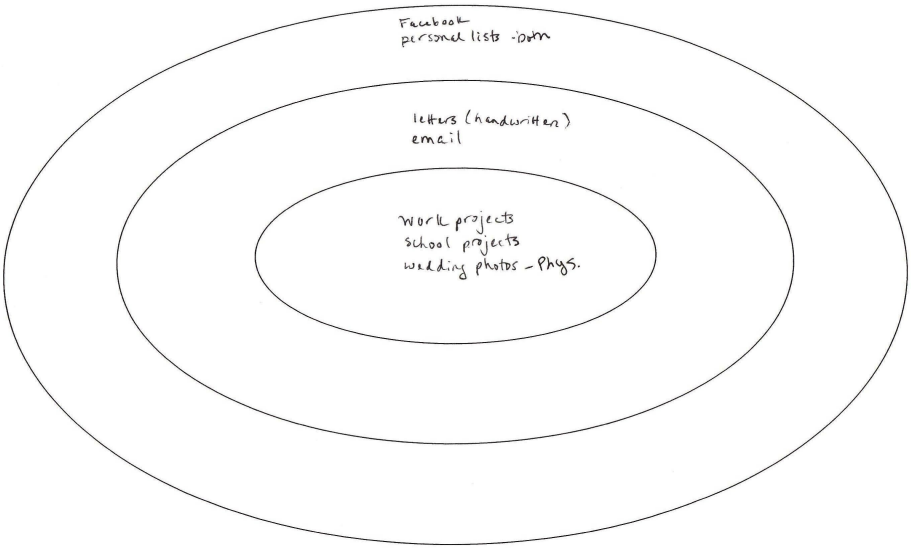


Information Source Horizon Maps (Zoey)

Everyday



Long Term



Appendix E: Content Stored in Digital Devices by Individual Participant

Participant	Number of	Number	Most Distributed Content Type and

	Devices	of Content Types	Number of Devices Used to Store that Content
Eleanor	6	12	PDFs, Spreadsheets & Web Pages - 4
Frank	8	14	Digital Video, Word Documents, PowerPoint, & Spreadsheets - 4
Greg	5	6	Photos – 5
Holly	6	8	Photos – 6
Irene	3	7	Photos – 3
Julie	10	14	Photos – 5
Kelly	9	12	Photos – 7
Lisa	6	10	Music & Photos – 5
Mark	9	15	Digital Video – 7
Nora	4	15	Digital Video – 4
Oscar	9	18	Database Files – 7
Paula	4	7	Music – 3
Quincy	3	8	Music, Photos & Spreadsheets - 2
Robert	7	11	Database, Photos, & PowerPoint - 5
Sarah	6	11	Photos – 6
Tina	7	13	Word Documents – 7
Uri	6	10	Digital Video & Photos 6
Victoria	6	9	Database Files– 5
Wendy	8	10	Photos – 6
Xavier	8	9	Photos – 8
Yvonne	7	8	Database & Photos – 4
Zoey	7	9	Digital Video & Photos – 5

Appendix F: Content Stored on the Web by Individual Participant

Participant	Web Locations	Content Types

Eleanor	2	6
Frank	5	11
Greg	2	2
Holly	4	5
Irene	2	2
Julie	1	2
Kelly	2	5
Lisa	2	2
Mark	5	9
Nora	2	10
Oscar	4	10
Paula	5	2
Quincy	3	1
Robert	3	3
Sarah	1	1
Tina	2	4
Uri	3	2
Victoria	0	0
Wendy	5	2
Xavier	4	6
Yvonne	2	2
Zoey	4	2

Appendix G: Content Selected for Redundant Storage by Individual Participant

	Do you treat	Do you make	If some, what content?

	content intended for long term use differently?	duplicate copies some of your content?	
Cathy	Yes	Some	Photos
David	No	Some	Photos
Eleanor	No	All	
Frank	No	All	
Greg	Yes	Some	Photos
Holly	No	Some	Contents of PDA
Irene	No	Some	Book Manuscript; Some Music and Photos
Julie	No	Some	Photos
Kelly	Yes	Some	Photos, School Papers
Lisa	No	All	
Mark	Yes	All	
Nora	Yes	Some	Current Work
Oscar	Yes	All	Back up Image Lap Top
Paula	No	Some	School Papers
Quincy	Yes	Some	Music
Robert	No	None	
Sarah	Yes	Some	Photos, School Papers
Tina	No	All	Back up Image Desk Top
Uri	No	Some	Photos
Victoria	No	All	
Wendy	Yes	Some	School Papers, Photos, Music
Xavier	Yes	Some	Photos, Creative Writing, Letters, Emails
Yvonne	Yes	All	Photos, documents
Zoey	Yes	Some	Creative Writing, Specific Emails

Appendix H: Where, How Often and by Whom Content is Duplicated by Individual Participant

	Where is the content duplicated?	When?	By Whom?
Cathy	Desk Top (DT) to External Hard Drive (EHD)	When appropriate	Husband
David	DT to Compact Disks (CD)	When Appropriate	Self
Eleanor	Lap Top (LT) to EHD, Personal Digital Assistant (PDA) to LT	Every two weeks	Husband
Frank	LT to EHD	Once a week	
	LT to Flash Drive (FD)	Daily	
	FD to Web Storage (WS)	Daily	Self
Greg	LT/DT to EHD	When appropriate	Self
Holly	PDA to LT	Once a week	Self
Irene	LT to Email; LT to CD	When appropriate	Self
Julie	LT/DT to FD & WS	When appropriate	Self
Kelly	LT to FD, Paper(P) & CD	When Appropriate	Self
Lisa	DT to EHD & WS	When I get worried enough about it	Self
Mark	DT to EHD, DVD & CD; EHD to EHD	Every couple of days	Self
Nora	DT to P & DT to CDROM	When Appropriate	Self and Husband
Oscar	LT/DT to EHD, DVD & WS	Daily	Self
Paula	LT/DT to P	When Appropriate	Self
Quincy	LT/DT to EHD & FD	When Appropriate	Self
Robert	Used to use flash drive, but lost it	Does not do so	
Sarah	DT/LT to Flash Drive & CD	When Appropriate	Self
Tina	LT to EHD and FD	Daily	Automatic
Uri	DT to DT & DVD	When Appropriate	Self
Victoria	DT to EHD	Every couple months	Self
Wendy	LT to EHD, FD & Email	When Appropriate	Self
Xavier	LT/DK to EHD, FD, DVD & WS	Every two weeks.; monthly for DVDs	Self
Yvonne	LT to EHD & FD	Every six months or so	Husband
Zoey	LT/DT to EHD& FD	Every six months or so	Self

Appendix I: Storage Locations of Most Valued Digital Information

Most Valued Digital Information located in Multi-functional digital devices

by individual and for all participants

	Desk Top	Lap Top	Cell Phone	PDA	iPod	
Photos	Lisa, Nora, Quincy, Robert, Uri*, Xavier	Frank, Irene, Quincy, Wendy, Xavier, Zoey	Nora			10
School work	Nora	Wendy, Yvonne				3
Personal Writing	Xavier	Wendy, Xavier, Zoey, Frank				4
Music	Nora, Xavier	Holly, Xavier			Holly, Nora	3
Email	Lisa, Sarah, Tina, Xavier					4
Calendar/planner	Julie, Robert, Tina			Julie		3
Work Documents		Yvonne				1
Will	Julie			Julie		1
Passwords		Oscar				1
Volunteer projects	Nora					1
Taxes	Sarah					1
Contact info			Xavier			1
Video		Oscar				1
Total Number of Participants	9	9	2	1	2	

* Stored on multiple desk stops.

Most Valued Digital Information located in storage devices individual and for all participants

	External Hard Drive	Flash Drive	CD	DVD	Floppy	
Photos	Frank, Lisa, Xavier, Zoey	Frank	Nora, Uri	Uri, Xavier		6
School work	Wendy	Wendy, Yvonne	Nora		Nora	3
Personal Writing	Wendy, Xavier, Zoey, Frank	Wendy, Frank		Xavier		4
Music	Xavier		Nora	Xavier		2
Email	Lisa, Tina, Xavier			Xavier		3
Calendar/planner	Tina					1
Work Documents	Mark*	Yvonne				2
Digital Art			Irene, Nora			2
Passwords	Oscar					1
Volunteer projects			Nora			1
Contact info				Xavier		1
Video	Oscar					1
Total Number of Participants	8	3	3	2	1	

* Stored on several external hard drives.

Most Valued Digital Information located on the Web by individual and for all participants

	Email Site	Subscription Web Storage	Kodak web site	Work Server	Total Number of Participants with Item Type
Photos		Frank	Robert		2
School work	Wendy, Yvonne				2
Personal writing/documents	Wendy	Frank			1
Calendar/planner				Julie	1
Work Documents	Yvonne				
Will				Julie	1
Passwords		Oscar			1
Total Number of Participants Using Location	2	2	1	1	

Most Valued Digital Information printed to paper by individual and for all participants

	Photos	School Work	Personal Writing	Calendar/ Planner	Will	Digital Art	Volunteer Projects	Taxes	Total Parti-cipants
Paper	Nora	Nora	Frank	Julie	Julie	Irene	Nora	Sarah	5

Appendix J: Specific Physical and digital item of great value, the value of those items, and the storage location/s for each digital item by individual participant

David	Physical	Coffee Table/ Piano Bench	Connection to Family	
	Digital	50 to 100 digital photos	Creative/Aesthetic, Tells a Story, Documents Life Events	CD, Kodak Site
Eleanor	Physical	Painting	Creative/Aesthetic, Connection to Family, Monetary	
	Digital	First Website Designed	Connection Self, Creative/Aesthetic	CD
Frank	Physical	Rare Books	Monetary, Historical, Tells a Story	
	Digital	Designed Template for Note Taking	Creative/Aesthetic	Lap Top, External Hard Drive, Web Storage
Greg	Physical	Picture of Daughter From 2 Years Ago	Connection to Family	
	Digital	Bank Password	Financial	Not Backed Up
Holly	Physical	A novel, *Clowns of God*	Connection to Self	
	Digital	Nothing		
Irene	Physical	Picture of Son Looking Very Happy	Connection to Family	
	Digital	My Own Novel	Creative/Aesthetic	Lap Top, Email
Julie	Physical	Pictures of Mother Who Died Recently	Connection to Family	
	Digital	Scanned pictures of mother who died recently set to music	Connection to Family, Creative/ Aesthetic, Practicality of Digital Information	Multiple DVDs
Kelly	Physical	college portfolio	Preference for Tangible, Creative/Aesthetic	
	Digital	photo story of friend who died young set to music	Connection to Friends, Creative/Aesthetic	CD, Flash Drive, Lap Top
Lisa	Physical	adoption documents from China and Cambodia	Connection to Family, Irreplaceable	
	Digital	picture of three kids at beach	Connection to Family	Desk Top, Paper
Mark	Physical	Bronze Art Piece	Historic, Monetary, Tells a Story	
	Digital	scanned photo of great grandfather family, only photo in existence of the family.	Connection to family	Desk Top, External Hard Drive
Nora	Physical	eulogy for my dad's	Connection to Family, Creative/ Aesthetic	
	Digital	instruction manual for tutors	Creative/Aesthetic, Saves Time	CD, Desk Top, Paper
Oscar	Physical	Passport	Connection to Self	
	Digital	Quicken File	Financial information, Historical	Web Subscription Storage, External Hard Drive

Paula	Physical	picture of family at mom's second wedding	Connection to family	
	Digital	art history final	Creative/Aesthetic, Future Reference, Practicality of Digital Information	Not Backed Up
Quincy	Physical	masters degree	Connection to Self, Documents Life Events	
	Digital	voicemail messages	Future Reference, Practicality of Digital Information	Not Backed Up
Robert	Physical	"We will change the world" punk rock poster	Connection to Self	
	Digital	photo of first child	Connection to Family	Lap Top, Kodak Site
Sarah	Physical	picture of grandmother	Connection to Family	
	Digital	family photo	Connection to Family, Documents Life Events	Flash Drive, CD
Tina	Physical	birth certificate	Saves Time	
	Digital	Quick Books	Supports Work, Financial	External Hard Drive
Uri	Physical	News article about a plane that crashed that he should have been on	Connection to Self	
	Digital	My photo library	Creative/Aesthetic	DVD, Desk Top
Victoria	Physical	Photos	Connection to Family	
	Digital	Blackberry- calendar and address book	Practicality of Digital Information, Communication	Backed up every Few Months to Desk Top
Wendy	Physical	grandparents wedding photo	Connection to Family, Irreplaceable	
	Digital	college papers	Creative/Aesthetic	Flash & External Hard Drive
Xavier	Physical	My first journal	Connection to Self	
	Digital	files on computer from high school, school work and creative writings	Connection to Self, Creative/Aesthetic	Not Backed Up
Yvonne	Physical	wedding album	Connection to Family, Sharing, Documents Life Events	
	Digital	proposal for thesis	Creative/Aesthetic, Future Reference, Practicality of Digital Information	Flash drive, Lap Top, Work Desk Top
Zoey	Physical	wedding & engagement rings	Connection to Family, Monetary	
	Digital	photo album senior year in college	Documents Life Events	Snapfish, Lap Top, Desk Top

Appendix K: Positive Experience with Digital Information by participant

	Specific Experience	Values
Ann	Creating a CD with Photos of Sister for party and to share with siblings	Creative/Aesthetic, Connection to Family, Sharing
Bill	Sharing Photos from a Charity Event	Sharing, Documents Life Events
Cathy	Sharing Photos of restricted archive in Rome	Sharing, Documents Life Events
David	Using email to store contact information	Practicality of Digital Information
Eleanor	First Time using Digital Camera	Enjoyment of Technology
Frank	Created Website related to personal interests	Creative/Aesthetic, Supports Work, Sharing, Communication
Greg	Picture of Wife on Vacation	Connection to Family, Documenting Life Events
Holly	Text Messages from daughter overseas after a bombing	Connection to Family
Irene	Digital Video of Son	Connection to Family
Julie	Sharing jokes and stories received from other friends through email	Sharing
Kelly	Photo Story of Class	Creative/Aesthetic
Lisa	Created a flyer to sell her house	Creative/Aesthetic, Supports Work
Mark	Sharing photos of nephew's wedding with mom through email	Connection to Family, Documents Life Events, Sharing
Nora	Receiving cartoon created by niece through email	Connection to Family
Oscar	First Party Play List Created using iTunes	Creative/Aesthetic, Enjoyment of Technology, Sharing
Paula	Prom pictures shared through email and MySpace	Connection to Friends, Sharing, Documents Life Events
Quincy	Created Website related to interests	Creative/Aesthetic, Supports Work
Robert	First Time Chatting with friends	Enjoyment of Technology, Sharing
Sarah	Sharing Photo of Classmates out for dinner	Connection to Friends, Sharing, Documents Life Events
Tina	Created Logo for Her Own Business	Creative/Aesthetic, Supports Work
Uri	Editing Photos to make them funny	Creative/Aesthetic
Victoria	Created a poem for Sister in Law	Connection to Family, Creative/Aesthetic
Wendy	Mom, Sister and Self sharing photos from their recent vacation	Connection to Family, Documents Life Events, Sharing
Xavier	Making Handwritten Poetry Digital	Sharing
Yvonne	Guests and self sharing digital photos after her wedding	Connection to Family and Friends, Sharing, Documents Life Events
Zoey	Creative writing stored on floppy discs that she had forgotten about	Creative/Aesthetic

Curriculum Vitae

Andrea Copeland Japzon

Education

Ph.D., Information Studies, September 2005 – June 2009
College of Information Science & Technology, Drexel University, Philadelphia, PA
Dissertation: Exploration of the motivation for and knowledge of digital preservation practices for personal digital information
Committee Chair: Denise E. Agosto

M.A., Geography, 2002
Department of Geography, Hunter College of the City University of New York, New York, NY
Thesis: A neighborhood analysis of public library use in New York City

M.S., Library Science, 1994
College of Information, Florida State University, Tallahassee, FL

B.A., English, 1992
Department of English, University of Florida, Gainesville, FL

Funded Awards

2005-2009, Laura Bush 21st Century Librarian Fellowship, Institute of Museum and Library Services.

2002-2003, SWG Fellowship, Society of Woman Geographers. Fellowship.

Selected Publications – Refereed Articles

Gong, H., Japzon, A., & Chen, C. (2008). Public libraries and social capital in three New York City neighborhoods. *Tijdschrift voor Economische en Sociale Geografie, 99*(1), 65-83.

Japzon, A. & Gong, H. (2005). A neighborhood analysis of public library use in New York City. *Library Quarterly, 45,* 446-463.

Professional Experience

1994 – 1999, Adult Services & Reference Librarian, The New York Public Library

1999 – 2003, Access Services Librarian, Hunter College Library, Hunter College of the City University of New York

2004 – 2005, Deputy Project Manager, Zimmerman Associates Inc., NASA Goddard Space Flight Center Library

(Blank)